Triggering on CP Violation

Aaij, R.J.M.; Merk, Marcel

2015

document version
Publisher's PDF, also known as Version of record

Triggering on *CP* Violation
Real-Time Selection and Reconstruction of
$B_s \to J/\psi\,\varphi$ Decays

R. Aaij

January 2015

Cover: Willem Aaij and Roel Aaij
Printed by: Proefschriftmaken.nl || Uitgeverij BOXPress
Published by: Uitgeverij BOXPress, 's-Hertogenbosch

This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM), which is part of The Netherlands Organisation for Scientific Research (NWO). It was carried out at the National Institute for Subatomic Physics (Nikhef) in Amsterdam, The Netherlands.

VRIJE UNIVERSITEIT

Triggering on *CP* Violation

Real-Time Selection and Reconstruction of $B_s \to J/\psi\,\varphi$ Decays

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Exacte Wetenschappen
op donderdag 7 mei 2015 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Roel Johannes Makis Aaij

geboren te Woerden

promotor:	prof.dr. H.G. Raven
copromotor:	prof.dr. M.H.M. Merk

Contents

Introduction **2**

1 Phenomenology of the $B_s \to J/\psi\varphi$ Decay **6**
1.1 Observable *CP*-violation . . . 6
1.1.1 *CP*-violation in Decay . . . 7
1.1.2 *CP*-violation in Mixing . . . 8
1.1.3 Time-Dependent *CP*-violation in the Interference between Mixing and Decay . . . 10
1.2 The Decay $B_s \to J/\psi\varphi$ . . . 11
1.2.1 Non–Resonant Decay of $B_s \to J/\psi K^+K^-$ . . . 13
1.2.2 *CP* Violation in $B_s \to J/\psi\varphi$ . . . 15
1.3 Higher Order Corrections (Penguins) . . . 16

2 The LHCb Detector at the LHC **18**
2.1 The Large Hadron Collider . . . 18
2.2 The LHCb Detector . . . 19
2.3 The Vertex Locator . . . 21
2.4 Silicon Tracker . . . 25
2.5 Outer Tracker . . . 26
2.6 Tracking . . . 27
2.7 The RICH Detectors . . . 29
2.8 The Calorimeters . . . 31
2.9 The Muon System . . . 35
2.10 The LHCb Trigger . . . 35
2.10.1 The Hardware Trigger . . . 37
2.10.2 High Level Trigger . . . 38

3 The LHCb High Level Trigger **40**
3.1 Measuring HLT Performance . . . 42
3.2 HLT1 . . . 44
3.2.1 Velo Track Finding . . . 46
3.2.2 Matching Velo Tracks to Muon Hits . . . 47
3.2.3 Primary Vertex Reconstruction . . . 58
3.2.4 Forward Track Finding . . . 58
3.2.5 Muon Identification . . . 60
3.2.6 Track Fitting . . . 60
3.2.7 HLT1 Dimuon Selections . . . 61
3.2.8 HLT1 Single Track Selections . . . 64
3.3 HLT2 . . . 68
3.3.1 Reconstruction in HLT2 . . . 69
3.3.2 HLT2 J/ψ Selections . . . 70
3.3.3 Inclusive HLT2 Selections . . . 71

4 Analysis **74**
4.1 Reconstruction and Selection of $B_s \to J/\psi\,\varphi$ Candidates . . . 74
4.1.1 Primary Vertex Assignment . . . 75
4.1.2 Decay-Time . . . 75
4.1.3 Mass PDF . . . 76
4.1.4 Selection of $B_s \to J/\psi\,\varphi$ Candidates . . . 78
4.1.5 Multiple Candidates . . . 79
4.1.6 Trigger Selections . . . 82
4.1.7 Reflection Backgrounds . . . 84
4.2 Statistical Background Subtraction . . . 84
4.3 S-Wave and K^+K^- Invariant Mass Distribution . . . 86
4.4 Decay-Time Acceptance . . . 88
4.5 Decay-Time Resolution . . . 90
4.5.1 Sources of Decay-Time Uncertainty . . . 90
4.5.2 Decay-Time Resolution Model . . . 92
4.5.3 Calibration . . . 93
4.5.4 Effect of Resolution on CP-Violating Observables . . . 93
4.6 Angular Acceptance . . . 95
4.6.1 Iterative Reweighing . . . 97
4.7 Flavour Tagging . . . 97
4.7.1 Opposite-Side Taggers . . . 99
4.7.2 Same-Side Taggers . . . 101

4.7.3 Tagging Performance . . . 102
4.7.4 Incorporating Tagging in the Signal PDF . . . 103
4.8 Observable PDF . . . 104

5 Decay-Time Acceptance 106
5.1 Decay-Time Acceptance from Reconstruction . . . 107
5.2 Fitting for the Decay-Time Acceptance . . . 110
5.3 Per-Event Decay-Time Acceptance (Swimming) . . . 112
5.3.1 Background and Justification . . . 114
5.3.2 Swimming Algorithm . . . 115
5.3.3 Swimming Validation . . . 117

6 Decay-Time Resolution 120
6.1 Calibration Procedure . . . 120
6.2 Calibration on Data . . . 122
6.2.1 Fit with Constant Scale-Factors . . . 123
6.2.2 Reparameterisation of the Resolution Model . . . 125
6.2.3 Calibration of the Per-Event Error Parameterisation . . . 126
6.3 Cross-checks with Simulated Candidates . . . 128
6.3.1 Fits to Simulated Samples . . . 131
6.3.2 Parameterisation of the Common Mean, μ . . . 133
6.3.3 Calibration of the Effect of Momentum . . . 138
6.3.4 Systematic Uncertainties . . . 138
6.4 Primary Vertex Refitting . . . 142
6.5 Wrong Primary Vertex Association . . . 146
6.5.1 Event Mixing Procedure . . . 146
6.5.2 Fits With Event Mixed Distributions . . . 147
6.5.3 Data . . . 149
6.5.4 Simulated Samples . . . 150
6.5.5 Signal Candidates with a Wrongly Assigned PV . . . 154
6.5.6 Future Considerations . . . 156

7 Results 158
7.1 Results . . . 158
7.2 Systematic Uncertainties . . . 160
7.2.1 Mass Model Uncertainties . . . 162
7.2.2 Reflection Backgrounds . . . 162
7.2.3 Factorisation . . . 164
7.2.4 K^+K^- mass integrals . . . 164

7.2.5 $B_c^+ \to B_s X$ . . . 165
7.2.6 Decay-Time Resolution . . . 165
7.2.7 Wrong Primary Vertex Assignment . . . 166
7.2.8 Decay-Time Acceptance . . . 166
7.2.9 Angular Resolution . . . 167
7.2.10 Angular Acceptance . . . 167
7.2.11 Correlations between Angular and Decay-Time Acceptances . . . 168
7.2.12 Discussion . . . 168
7.3 Conclusions . . . 169

A Streamer Framework **172**

B Numerical Calculation of the Dilution **174**

C Effect of Using sWeights for Resolution Fits **176**

D Reparameterisation of a Double Gaussian **180**

Bibliography **184**

Summary **194**

Samenvatting **198**

Dankwoord **204**

Introduction

With the realisation of the theories of mechanics, electrodynamics and thermodynamics at the end of the 19th century, physics was believed to be nearly complete. Few things remained unexplained and many physicists felt that nature was fully understood. Everything changed in 1905 with Einstein's realisation that the theory of electrodynamics did not fit in the framework of classical dynamics, and the creation of his theory of special relativity [1]. The development of quantum physics during the 1920s brought further upheaval and it became clear that nature was far from understood.

In 1928, while creating a relativistic theory of quantum mechanics, Paul Dirac proposed the existence of so-called anti-particles [2]. When the anti-particle of the electron, the positron, was discovered in 1932 [3], the existence of anti-particles was confirmed and a new field of physics was born: particle physics. Since the positron was indeed found to be equal but opposite to the electron in all aspects except mass and spin, as predicted by Dirac, nature was assumed to be symmetric between particles and anti-particles. This symmetry implied that if sufficient energy was available, anti-particles and particles could only be created in equal amounts. If they came into contact again, they would annihilate and only light would remained.

Around the same time that anti-particles were first proposed and discovered, observations of distant galaxies by Lemaître [4] and Hubble [5] showed that the universe was expanding. Assuming constant expansion, it was concluded to have been infinitely small approximately 13.7 billion years ago. Thus the idea of the Big Bang arose. In the first instants after the Big Bang, the temperature of the universe was so high, that matter could not exist. The matter we observe in the universe today must therefore have been created at a later time.

Large scale observations of the universe and the AMS experiment [6] have since confirmed that at the present time, there are only very small amounts of anti-matter in the universe. This implies that at some point, through some mech-

anism, anti-matter must have disappeared, while matter remained. While three requirements for the generation of a matter-anti-matter asymmetry were identified by Sakharov [7], the exact mechanism that generates the asymmetry is unclear at this time.

Developments in experimental particle physics since the 1930s have resulted in the discovery of many more elementary particles. Inspired by these discoveries, theoretical developments have resulted in the creation of a Standard Model that describes all of these particles and their weak, strong and electromagnetic interactions, whereas gravitational interactions are not included. Five additional particles similar to the electron have been discovered, the muon, the tau and three neutrinos. Such particles are called leptons, and do not take part in the strong interaction. Six particles that do interact strongly, called quarks, were discovered; down, up, strange, charm, beauty and top. The interactions between these particles are mediated by bosons: the photon for the electromagnetic interaction, the $W^{\pm}$ and Z for the weak interaction and the gluon for the strong interaction. The success of the Standard Model culminated in 2013 with the discovery [8, 9] of the last remaining particle whose existence was predicted [10, 11] by the Standard Model, but had not yet been discovered: the Higgs boson.

Due to the nature of the strong interaction, quarks are only found in bound states of two or three, respectively called mesons and baryons. Mesons that consist of a beauty quark and another quark are called B mesons. In the Standard Model, B mesons that carry no electrical charge can oscillate into their anti-particles and vice-versa. For B^0 mesons, which consist of a b and a d quark, half of an oscillation period is approximately equal to two-thirds the average lifetime of the meson, while for B_s mesons half a period only takes approximately 2% of the average lifetime of the meson. These fast oscillations provide an opportunity for studying differences between matter and anti-matter.

One of the requirements for the generation of a matter-anti-matter symmetry identified by Sakharov is the violation of the matter-anti-matter symmetry, *C*, and its combination with the left-right symmetry, *CP*. While the weak interaction maximally violates *P* [12], *CP*-violation is much smaller in the Standard Model; too small, in fact, to generate the observed matter-anti-matter asymmetry of the universe. Hence there must be physics beyond the Standard Model.

To study differences between matter and anti-matter, decays of $B_s \to J/\psi\,\varphi$ are particularly interesting. For this decay, the predicted amount of *CP*-violation, expressed by the parameter ϕ_s, is very small [13], and any deviations from the prediction are a clear sign of physics beyond the Standard Model. In addition, there are several experimental advantages to the study of this decay: the branch-

ing fraction is $O(10^{-5})$, the presence of muons in the final state allows for a straightforward trigger with good efficiency, and there is little background. The precise theoretical prediction and experimental advantages of the $B_s \to J/\psi\varphi$ decay mode resulted in it being labelled a "golden" mode for LHCb.

The measurement of ϕ_s with 3 fb^{-1} of data collected in 2011 and 2012 by the LHCb experiment is the goal of the analysis described in this thesis. Chapter 1 gives a general introduction on the measurement of *CP*-violating observables, it describes the phenomenology of the $B_s \to J/\psi\varphi$ decay, and it introduces the decay-rate equations. The LHCb experiment is described in chapter 2, followed by a discussion of its High Level Trigger in chapter 3. The $B_s \to J/\psi\varphi$ analysis is described in chapter 4, followed by a discussion of the decay-time dependent efficiency model in chapter 5, and the calibration of the decay-time resolution model in chapter 6, before the results of the analysis are presented in chapter 7.

Chapter 1

Phenomenology of the $B_s \to J/\psi\,\varphi$ Decay

The decay of $B_s \to J/\psi\,\varphi$ is interesting to search for *CP*-violation, due to a theoretically clean prediction of small *CP*-violation in the Standard Model (SM). This chapter describes the requirements for observing *CP*-violation in general in section 1.1, and for $B_s \to J/\psi\,\varphi$ decays in particular in section 1.2. The physics parameters of interest are introduced, and their Standard Model predictions are given. These parameters will be determined with a likelihood fitting procedure matching a theoretical probability distribution function (PDF) to the data. The PDF that describes the decay rates of $B_s \to J/\psi\,\varphi$ decays is introduced in section 1.2, which forms the basis of the measurement of ϕ_s described in this thesis.

1.1 Observable *CP*-violation

CP Violation occurs as the result of interfering quantum amplitudes with *CP*-violating phase differences. These phase differences can be generated through the introduction of complex-valued coupling constants in the theory. In the case of the strong interactions, this would however lead to an electric dipole moment of the neutron d_N [14, Sec 9.1], and [15], which is currently excluded at $d_N < 2.9 \times 10^{-26}$. This leads to a limit on the *CP*-violating phase in the strong interaction, $\phi < 10^{-10}$ [16]. In the weak sector of the SM Lagrangian, *CP*-violation originates from the single complex degree of freedom in the unitary CKM matrix [17, 18], which in turn leads to the presence of several complex phases in flavour changing current interactions, as described in the CKM model.

Since physical observables that originate from a single amplitude are invariant under global phase transformations, the observation of any complex phase (difference) requires the coherent contribution of at least two interfering amplitudes, for example A_1 and A_2. If the phase difference between these two amplitudes is assumed to consist of two contributions, one that changes sign under *CP*, ϕ, and one that does not, δ, the amplitudes can be written as:

$$A_1 = |A_1|, \quad A_2 = |A_2|e^{i\delta}e^{i\phi}, \tag{1.1}$$

where the phase of amplitude A_1 has been chosen to be equal to zero. The decay-rate of an unstable particle as a result of these amplitudes becomes:

$$\begin{aligned} |A|^2 &= |A_1 + A_2|^2 \\ &= ||A_1| + |A_2|e^{i\delta}e^{i\phi}|^2 \\ &= |A_1|^2 + |A_2|^2 + 2|A_1||A_2|\cos(\delta + \phi), \end{aligned} \tag{1.2}$$

while the decay rate of its *CP*-conjugate process becomes:

$$\begin{aligned} |\bar{A}|^2 &= |A_1 + A_2|^2 \\ &= ||A_1| + |A_2|e^{i\delta}e^{-i\phi}|^2 \\ &= |A_1|^2 + |A_2|^2 + 2|A_1||A_2|cos(\delta - \phi). \end{aligned} \tag{1.3}$$

The *CP*-asymmetry, A_{CP}, is then equal to:

$$A_{CP} = \frac{|A|^2 - |\bar{A}|^2}{|A|^2 + |\bar{A}|^2} = \frac{-2\sin\phi\sin\delta}{r + 1/r + 2\cos\phi\sin\delta}, \tag{1.4}$$

where $r = |A_1|/|A_2|$.

From this, it can be seen that an observable difference between the decay-rate as a consequence of *CP*-violation can only occur if both $\sin\delta \neq 0$, and $\sin\phi \neq 0$. Furthermore, the observable asymmetry is largest if $r = 1$, i.e. $|A_1| = |A_2|$, and vanishes if either amplitude vanishes. Observable *CP*-violating processes are generally categorised in three categories, which are discussed in the following sections.

1.1.1 *CP*-violation in Decay

CP-violation in decay occurs if the magnitude of the (total) amplitude, A_f for a B meson to decay to final state f, is not equal to the magnitude of the (total)

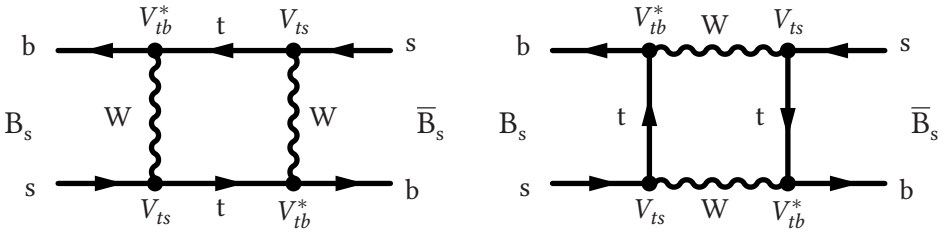


Figure 1.1: Feynman box diagrams for B_s meson mixing. The charge-conjugated process for $\overline{B}_s \to B_s$ is obtained by replacing all quarks by anti quarks (and vice versa) and taking the complex-conjugates of the CKM elements.

amplitude, $\bar{A}_{\bar{f}}$, of the *CP*-conjugate process:

$$\left|\frac{A_f}{\bar{A}_{\bar{f}}}\right| \neq 1. \tag{1.5}$$

This, for example, is the case if the phase difference that changes sign under *CP* arises from the differences between contribution to the decay amplitude of tree level Feynman diagrams and that of diagrams containing loops. The phase difference that does not change sign under *CP* arises from the strong interactions between the final state particles. *CP*-violation in decay was first discovered in kaons [19], and is now also observed in B-decays, for example, in decays of $B^0 \to K^+\pi^-$ and $B_s \to K^-\pi^+$ [20].

1.1.2 *CP*-violation in Mixing

Due to the presence of the box diagrams shown in fig. 1.1, and the fact that the mass and flavour eigenstates of B_s mesons are different, B_s mesons may oscillate into $\overline{B}_s$ mesons, and vice-versa. The mass eigenstates can be written as a superposition of the flavour eigenstates as:

$$\begin{aligned} |B_{s,H}\rangle &= p\,|B_s\rangle - q\,|\overline{B}_s\rangle\,, \\ |B_{s,L}\rangle &= p\,|B_s\rangle + q\,|\overline{B}_s\rangle\,, \end{aligned} \tag{1.6}$$

with $|p|^2 + |q|^2 \equiv 1$. The mass difference, Δm_s, and decay-width difference, $\Delta\Gamma_s$, are then defined as:

$$\begin{aligned} \Delta m_s &= m_H - m_L, \\ \Delta\Gamma_s &= \Gamma_H - \Gamma_L. \end{aligned} \tag{1.7}$$

By solving the Schrödinger equation for such an unstable two-state system following the procedure described in [21, 22] for a system of atomic energy transitions, and transforming the solution to the flavour state basis, the time evolution of initially pure B_s and $\overline{B}_s$ mesons is given by [23, 24, 25]:

$$\begin{aligned} |B_s(t)\rangle &= g_+(t)\,|B_s\rangle + \frac{q}{p} g_-(t)\,|\overline{B}_s\rangle \\ |\overline{B}_s(t)\rangle &= \frac{p}{q} g_-(t)\,|B_s\rangle + g_+(t)\,|\overline{B}_s\rangle\,, \\ g_\pm(t) &= \frac{1}{2}\left(e^{-(im_L+\Gamma_L/2)t} \pm e^{-(im_H+\Gamma_H/2)t}\right). \end{aligned} \tag{1.8}$$

The probabilities for an initially produced B_s to decay as $\overline{B}_s$, and vice-versa, are then given by:

$$\begin{aligned} |\langle B_s(t)|\overline{B}_s\rangle|^2 &= \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \\ |\langle \overline{B}_s(t)|B_s\rangle|^2 &= \left|\frac{p}{q}\right|^2 |g_-(t)|^2 \end{aligned} \tag{1.9}$$

If these probabilities are not equal, in other words:

$$\left|\frac{q}{p}\right|^2 \neq \left|\frac{p}{q}\right|^2, \tag{1.10}$$

the rates of decays of initially produced B_q^0 and $\overline{B}_q^0$ mesons to a flavour specific state will be different. If semileptonic decays of B^0 and B_s mesons are chosen as the flavour specific final state, the corresponding *CP*-asymmetries are labelled a_{sl}^d and a_{sl}^s, and given by:

$$a_{sl}^q = \frac{\Gamma(\overline{B}_q^0 \to D_q^+ \mu^- \overline{\nu}_\mu X) - \Gamma(B_q^0 \to D_q^- \mu^+ \nu_\mu X)}{\Gamma(\overline{B}_q^0 \to D_q^+ \mu^- \overline{\nu}_\mu X) + \Gamma(B_q^0 \to D_q^- \mu^+ \nu_\mu X)}, \tag{1.11}$$

where the q may be d or s, and the muon is chosen, as it is the lepton that is easiest to detect experimentally . These asymmetries have recently been measured by LHCb to be $a_{sl}^s = -0.06 \pm 0.50 \pm 0.36\%$ [26] and $a_{sl}^d = -0.02 \pm 0.19 \pm 0.30\%$ [27], where the first uncertainties are statistical and the second systematic. These measurements are consistent with $|q/p| = 1$, i.e. no *CP*-violation in mixing.

1.1.3 Time-Dependent *CP*-violation in the Interference between Mixing and Decay

For final states that both a B_s and a $\overline{B}_s$ can decay to, an initially produced B_s might decay directly to such a final state, or first mix to a $\overline{B}_s$ before decaying, and vice-versa. In such decays, possibilities arise for *CP*-violation in the interference between mixing and decay, if a relative phase difference, that changes sign under *CP*, is present between the two decay paths. The *CP* even phase difference arises from the occurrence of g_+ and g_- — defined in eq. (1.8) — in the mixing process. The 90° phase difference becomes explicit when $\Delta\Gamma_s$ is set to zero, and the equations for $g_\pm$ are rewritten as:

$$\begin{aligned} g_+(t) &= e^{-im_s t} e^{-\Gamma_s t/2} \cos\frac{\Delta m_s t}{2}, \\ g_-(t) &= e^{-im_s t} e^{-\Gamma_s t/2}\, i \sin\frac{\Delta m_s t}{2}, \end{aligned} \tag{1.12}$$

where m_s and Γ_s are the average mass and decay width, respectively. To describe such *CP*-violation, the parameter λ is introduced, which is defined as:

$$\lambda_f \equiv \frac{q\bar{A}_f}{pA_f}. \tag{1.13}$$

If a final state is additionally a *CP* eigenstate, the following holds:

$$CP\,|f_{CP}\rangle = \eta_f\,|f_{CP}\rangle\,, \tag{1.14}$$

from which if follows that:

$$\bar{A}_f = \eta_f \bar{A}_{\bar{f}} \tag{1.15}$$

$$\lambda_f = \eta_f \frac{q\bar{A}_{\bar{f}}}{pA_f}. \tag{1.16}$$

Since an initially produced B meson can either directly decay to final state f_{CP} with an amplitude of $A_f g_+(t)$ or first mix and subsequently decay with an amplitude $\lambda_f A_f g_-(t)$, the total time dependent decay amplitude, T becomes:

$$\begin{aligned} T &= \langle f_{CP}|H|B(t)\rangle, \\ &= A_f[g_+(t) + \lambda_f g_-(t)], \end{aligned} \tag{1.17}$$

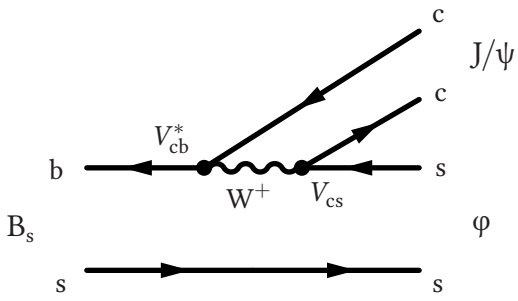


Figure 1.2: Feynman diagrams for the decay of $B_s \to J/\psi\,\varphi$.

where H is the Hamiltonian that describes the mixing and decay. The decay rates for a B_s decay to a *CP* eigenstate are obtained from eq. (1.17) by substituting the expressions for $g_\pm(t)$ given in eq. (1.8) and squaring:

$$\Gamma_{B_s\to f}(t) = \left|A_f\right|^2 (1+|\lambda_f|^2)\frac{e^{-\Gamma_s t}}{2}\times \left(\cosh\frac{\Delta\Gamma_s t}{2} + C\cos(\Delta m_s t) + D\sinh\frac{\Delta\Gamma_s t}{2} + S\sin(\Delta m_s t)\right), \tag{1.18}$$

where the following observables are introduced:

$$D = \frac{2\Re(\lambda_f)}{1+|\lambda_f|^2}, \quad C = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \quad S = \frac{-2\Im(\lambda_f)}{1+|\lambda_f|^2}, \tag{1.19}$$

with $C^2+S^2+D^2=1$. The rate for $\overline{B}_s$ decays is obtained applying the following substitutions to eq. (1.18): $C \to -C$, $S \to -S$, and by including a relative factor $|p/q|$. *CP*-violation occurs if the time-dependent rates of $B \to f$ and $\overline{B} \to f$ are not the same, for example if $S \neq 0$ or $C \neq 0$.

1.2 The Decay $B_s \to J/\psi\,\varphi$

A B_s decay to a *CP* eigenstate of particular interest is the decay $B_s \to J/\psi\,\varphi$. The tree-level Feynman diagrams for this decay are shown in fig. 1.2. The *CP* eigenvalues of the $J/\psi\,\varphi$ final state are $\eta_f = \pm 1$, depending on the total angular momentum, L, of the final state:

$$CP|J/\psi\,\varphi\rangle = \eta_f|J/\psi\,\varphi\rangle = (-1)^L|J/\psi\,\varphi\rangle. \tag{1.20}$$

The B_s is a spinless particle and therefore has no total angular momentum: $\mathcal{J} = 0$. Both the J/ψ and φ resonances are however vector mesons with spin 1. Adding the spin of the the J/ψ and φ results in a total spin of $S \in \{0, 1, 2\}$, and conservation of angular momentum thus requires that the total spin of the J/ψ and the φ is compensated by their orbital angular momentum, such that $L + S = 0$. The orbital angular momentum of the J/ψ and φ can therefore have values of $L \in \{0, 1, 2\}$. The states with $L \in \{0, 2\}$ are *CP* even ($\eta_f = 1$), while the state with $L = 1$ is *CP* odd ($\eta_f =$ -1).

Due to the appearance of η_f in eq. (1.17), the coefficient S that appears in front of the oscillating term in eq. (1.18), has a different sign for *CP* even and *CP* odd final states. If the angular momentum of the final state in $B_s \to J/\psi\,\varphi$ decays is not measured, the amplitude of the oscillations will be diluted, depending on the relative fraction of the *CP* odd and *CP* even state.

To statistically disentangle the mixture of $J/\psi\,\varphi$ states with orbital angular momentum $L \in \{0, 1, 2\}$, an angular analysis is performed. Instead of using the states with $L \in \{0, 1, 2\}$ as a basis, three independent linear polarisations are used:

- in the so-called parallel ($\parallel$) state, the polarisation vectors of the J/ψ and φ are both oriented parallel with respect to the direction of their momentum in the rest frame of the B_s;
- in the so-called perpendicular ($\perp$) state, the polarisation vectors of the J/ψ and φ are both perpendicular with respect to the direction of their momentum in the rest frame of the B_s, and parallel to each other;
- in the so-called zero (0) state, the polarisation vectors of the J/ψ and φ are both perpendicular with respect to the direction of their momentum in the rest frame of the B_s, and perpendicular to each other.

The $\parallel$ and 0 states are *CP* even, while the $\perp$ state is *CP* odd [24, p. 79].

In the presence of four final state particles, three angles are sufficient to describe the decay topology. The so-called helicity angles are adopted, which are defined as shown in fig. 1.3, and derived in [25, App. A]. By including the angular dependence, the time-dependent differential decay-rates for initial B_s mesons (eq. (1.18)) becomes a sum over six terms,

$$\frac{\mathrm{d}^4\Gamma(B_s \to f)}{\mathrm{d}\boldsymbol{\Omega}\mathrm{d}t} \sim \sum_{i=1}^{6} T_i(t) f_i(\boldsymbol{\Omega}), \tag{1.21}$$

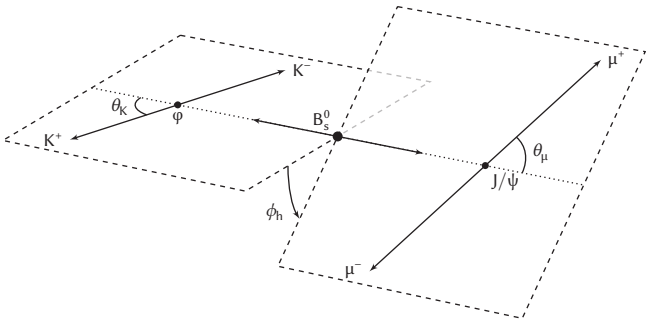


Figure 1.3: Definition of the so called helicity angles, θ_K, θ_μ, and ϕ_h; the momentum vectors are given in the rest frame of the respective mother particles. Figure kindly provided by [28].

where:

$$T_i(t) = N_i e^{-\Gamma_s t}\left(a_i \cosh\frac{\Delta\Gamma_s t}{2} + b_i \cos(\Delta m_s t) + c_i \sinh\frac{\Delta\Gamma_s t}{2} + d_i \sin(\Delta m_s t)\right). \tag{1.22}$$

The coefficients N_i are expressed in terms of the angular coefficients $A_i(t), i \in \{\|, \perp, 0\}$ at $t = 0$, A_i. The angular coefficients are parameterised by their magnitudes and phases as:

$$A_i = |A_i| e^{i\delta_i}, \tag{1.23}$$

where δ_0 is chosen to be zero, and $|A_\||^2 + |A_\perp|^2 + |A_0|^2 = 1$. The coefficients N_i, a_i, b_i, c_i, d_i, and the angular functions $f_i(\mathbf{\Omega})$ are given in table 1.1 for all terms, where the sign of a_i, b_i, c_i, and d_i as a result of the *CP* eigenvalue, η_f, is explicitly given. The differential decay rate for initially produce $\overline{B}_s$ mesons is obtained by applying the following substitutions: $b_i \to -b_i$, $d_i \to -d_i$, and by including a relative factor $|p/q|$, as before.

The complete and normalised probability distribution function (PDF) of the signal decay rate is then given by:

$$\mathrm{PDF}_{B_s \to J/\psi\,\varphi}(t, \mathbf{\Omega}) = \frac{\sum_{i=1}^{N} T_i(t) f_i(\mathbf{\Omega})}{\iint \mathrm{d}t\mathrm{d}\mathbf{\Omega} \sum_{i=1}^{N} T_i(t) f_i(\mathbf{\Omega})}, \tag{1.24}$$

1.2.1 Non–Resonant Decay of $B_s \to J/\psi K^+K^-$

The B_s can decay to the observed final state of $\mu^+\mu^-K^+K^-$ through a non-resonant K^+K^- state in addition to the $\varphi \to K^+K^-$ transition. The $J/\psi K^+K^-$ non-resonant

i	N_i	a_i	b_i	c_i	d_i	$f_i(\theta_\mu, \theta_K, \phi_h)$
1	$\lvert A_0\rvert^2$	1	C	D	$-S$	$2\cos^2\theta_K \sin^2\theta_\mu$
2	$\lvert A_\parallel\rvert^2$	1	C	D	$-S$	$\sin^2\theta_K \left(1 - \sin^2\theta_\mu \cos^2\phi_h\right)$
3	$\lvert A_\perp\rvert^2$	1	C	$-D$	S	$\sin^2\theta_K \left(1 - \sin^2\theta_\mu \sin^2\phi_h\right)$
4	$\lvert A_\parallel A_\perp\rvert$	$C\sin(\delta_\perp - \delta_\parallel)$	$\sin(\delta_\perp - \delta_\parallel)$	$S\cos(\delta_\perp - \delta_\parallel)$	$D\cos(\delta_\perp - \delta_\parallel)$	$\sin^2\theta_K \sin^2\theta_\mu \sin 2\phi_h$
5	$\lvert A_0 A_\parallel\rvert$	$\cos(\delta_\parallel - \delta_0)$	$C\cos(\delta_\parallel - \delta_0)$	$D\cos(\delta_\parallel - \delta_0)$	$-S\cos(\delta_\parallel - \delta_0)$	$\frac{1}{2}\sqrt{2}\sin 2\theta_K \sin 2\theta_\mu \cos\phi_h$
6	$\lvert A_0 A_\perp\rvert$	$C\sin(\delta_\perp - \delta_0)$	$\sin(\delta_\perp - \delta_0)$	$S\cos(\delta_\perp - \delta_0)$	$D\cos(\delta_\perp - \delta_0)$	$-\frac{1}{2}\sqrt{2}\sin 2\theta_K \sin 2\theta_\mu \sin\phi_h$

Table 1.1: Angular and time-dependent components of the P-wave component of the $B_s \to J/\psi\,\varphi$ signal PDF. Differences between phases that are invariant under *CP* are given as the sine or cosine of differences $\delta_i - \delta_j$, where $i, j \in \{\parallel, \perp, 0\}$ and $i \neq j$, as appropriate.

i	N_i	a_i	b_i	c_i	d_i	$f_i(\theta_\mu, \theta_K, \phi_h)$
7	$\lvert A_S\rvert^2$	1	C	$-D$	S	$\frac{2}{3}\sin^2\theta_\mu$
8	$\lvert A_S A_\parallel\rvert$	$C\cos(\delta_\parallel - \delta_S)$	$\cos(\delta_\parallel - \delta_S)$	$S\sin(\delta_\parallel - \delta_S)$	$D\sin(\delta_\parallel - \delta_S)$	$\frac{1}{3}\sqrt{6}\sin\theta_K \sin 2\theta_\mu \cos\phi_h$
9	$\lvert A_S A_\perp\rvert$	$\sin(\delta_\perp - \delta_S)$	$C\sin(\delta_\perp - \delta_S)$	$-D\sin(\delta_\perp - \delta_S)$	$S\sin(\delta_\perp - \delta_S)$	$-\frac{1}{3}\sqrt{6}\sin\theta_K \sin 2\theta_\mu \sin\phi_h$
10	$\lvert A_S A_0\rvert$	$C\cos(\delta_0 - \delta_S)$	$\cos(\delta_0 - \delta_S)$	$S\sin(\delta_0 - \delta_S)$	$D\sin(\delta_0 - \delta_S)$	$\frac{4}{3}\sqrt{3}\cos\theta_K \sin^2\theta_\mu$

Table 1.2: Angular and time-dependent components of the S-wave component of the $B_s \to J/\psi\,\varphi$ signal PDF.

final state is an S-wave, which makes it a purely *CP* odd final state. To take this contribution into account in the decay rate, the sum over all combinations polarisation states of $\{\|, \perp, 0\}$ given in eq. (1.21) is extended to include combinations with the polarisation of the non-resonant S-wave, which has amplitude A_S. The additional four terms are given in table 1.2.

To ensure that relative normalisation of the polarisation states is unchanged, the fraction of the non-resonant component, F_S, is defined as:

$$F_S = \frac{|A_S|^2}{|A_S|^2 + 1}. \tag{1.25}$$

A contribution from the alternative non-resonant decay of $B_s \to \varphi\mu^+\mu^-$ is also present, which would result in the appearance of a number of additional terms in the decay rate equation, but given the small branching fraction of $B_s \to \varphi\mu^+\mu^-$ relative to the branching fraction of $B_s \to J/\psi\,\varphi$ of: $\mathcal{B}(J/\psi\,\varphi)/\mathcal{B}(\varphi\mu^+\mu^-) = (6.74^{+0.61}_{-0.56} \pm 0.16) \times 10^{-4}$[29], this contribution is neglected.

1.2.2 *CP* Violation in $B_s \to J/\psi\,\varphi$

For $B_s \to J/\psi\,\varphi$ decays, one can define the convention dependent[1] phases of the mixing amplitude, ϕ_M and of the $b \to c\bar{c}s$ transition, ϕ_D according to [24, p. 23,88]:

$$\begin{aligned} \frac{q}{p} &= e^{-i\phi_M}[1 - a^s_{sl}], \\ \frac{\bar{A}_{J/\psi\,\varphi}}{A_{J/\psi\,\varphi}} &= -\eta_{J/\psi\,\varphi} e^{2i\phi_D}. \end{aligned} \tag{1.26}$$

It follows that:

$$\begin{aligned} \lambda_{J/\psi\,\varphi} &= \frac{q\bar{A}_{J/\psi\,\varphi}}{pA_{J/\psi\,\varphi}} \\ &= -\eta_{J/\psi\,\varphi} e^{-i\phi_M} e^{2i\phi_D}[1 - a^s_{sl}] \\ &= -\eta_{J/\psi\,\varphi} e^{-i\phi_s}[1 - a^s_{sl}], \end{aligned} \tag{1.27}$$

where $\phi_s = \phi_M - 2\phi_D$ is convention independent. In addition, from the diagrams in fig. 1.1 and fig. 1.2, it can be seen that in the Standard Model the convention dependent phases are given by:

$$\begin{aligned} \phi_M &= \arg[(V_{tb}V^*_{ts})^2] \\ \phi_D &= \arg(V_{cb}V^*_{cs}), \end{aligned} \tag{1.28}$$

[1] Here the same phase convention as in the Wolfenstein parameterisation is used [30].

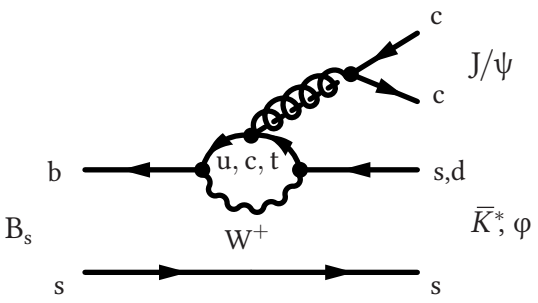


Figure 1.4: Example of a penguin diagram for the decays of $B_s \to J/\psi\,\varphi$ and $B_s \to J/\psi\,K^*$.

from which it follows that the convention independent phase difference is given by:

$$\begin{aligned}\phi_s &= \phi_M - 2\phi_D \\ &= \arg[(V_{tb}V_{ts}^*)^2] - 2\arg(V_{cb}V_{cs}^*) \\ &= 2\arg(\frac{V_{tb}V_{ts}^*}{V_{cb}V_{cs}^*}), \\ &= -2\beta_s, \end{aligned} \tag{1.29}$$

where β_s is an angle of one of the so-called Unitarity Triangles [14, Sec. 12.1]. In the Standard Model, β_s is predicted to be very small: $-2\beta_s$ = -0.0363$^{+0.0014}_{-0.0012}$ rad [13]. As a result of this prediction, a deviation of the measured value of ϕ_s from -2β_s would be a clear sign of physics beyond the Standard Model.

1.3 Higher Order Corrections (Penguins)

In the derivation of the value of ϕ_s in the Standard Model in eq. (1.29), only the tree-level diagram shown in fig. 1.2 was considered, and higher order loop-level diagrams were neglected. This class of diagrams is also referred to as penguin diagrams[31]. An example of such a diagram for the decay of $B_s \to J/\psi\,\varphi$ is shown in fig. 1.4.

Including penguin amplitudes, the $\bar{b} \to \bar{c}c\bar{s}$ decay amplitudes becomes:

$$\begin{aligned} A(\bar{b} \to \bar{c}c\bar{s}) &= V_{cs}V_{cb}^*(T_c + P_c) + V_{us}V_{ub}^*P_u + V_{ts}V_{tb}^*P_t \\ &= V_{cs}V_{cb}^*(T_c + P_c - P_t) + V_{us}V_{ub}^*(P_u - P_t), \end{aligned} \tag{1.30}$$

where, after factoring out the CKM elements, T_c is the tree-level amplitude, and P_i are the penguin amplitudes with $i \in \{u, c, t\}$. Unitarity of the CKM matrix was assumed in the second reordering step of eq. (1.30), such that $V_{ts}V_{tb}^* = -V_{us}V_{ub}^* - V_{cs}V_{cb}^*$ holds. The second term in the right-hand side of eq. (1.30) is much smaller than than the first term, since $|V_{us}V_{ub}^*| \approx \lambda^4$ versus $|V_{cs}V_{cb}^*| \approx \lambda^2$, where λ is one of the parameters of the Wolfenstein parameterisation of the CKM matrix[30]: $\lambda = |V_{us}| \approx 0.22$[13]. In addition to this suppression, only differences of penguin amplitudes appear in eq. (1.30), which further suppresses contribution to the decay amplitude relative to the tree-level amplitude.

The size of the penguin diagram contribution is hard to calculate from QCD, but might result in a value for $\sin\phi_s$ as large as -0.1[32], significantly larger than the predicted value of $\sin\phi_s = -0.03$ in case penguin contributions are ignored. It was suggested in Ref. [32] to exploit s–d flavour symmetry to measure the size of the penguin contributions from the alternative decay of $B_s \to J/\psi K^*$; the corresponding Feynman diagram is included in fig. 1.4. For this decay, the decay amplitude is defined analogously to eq. (1.30), where all s indices are replaced with d:

$$A(\bar{b} \to \bar{c}c\bar{s}) = V_{cd}V_{cb}^*(T_c + P_c - P_t) + V_{ud}V_{ub}^*(P_u - P_t). \tag{1.31}$$

In this case, two terms on the right hand side contribute equally to the amplitude, since they are both of order λ^3, which allows a measurement of the penguin contributions.

The suppression of the penguin amplitudes relative to the tree-level amplitudes in $B_s \to J/\psi\,\varphi$ is the main reason that the uncertainties on the theoretical prediction of ϕ_s in the Standard Model are as small as they are. Once the experimental uncertainty approaches the uncertainty on the theoretical prediction, the penguin contribution can be measured using decays of $B_s \to J/\psi K^*$. Currently this work is progressing in LHCb and results on the combined 2011 and 2012 data sets are expected for the summer of 2015.

As a starting point of the description of the measurement of ϕ_s from $B_s \to J/\psi\,\varphi$ decays, the LHCb detector is described in the following chapter.

Chapter 2

The LHCb Detector at the LHC

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is currently the world's largest and highest energy particle collider. It is a circular collider with a circumference of approximately 27 kilometres, located 100 metres underground near Geneva, Switzerland. Two bunched beams of protons circulate in opposite directions, and collide at four so-called interaction points. The LHCb detector is installed at one of these interaction points, Point 8, located near Geneva airport.

The main LHC ring is installed in the tunnel that was originally excavated to house the Large Electron Positron Collider (LEP), which operated from 1989 until 2000. Due to the low level of synchrotron radiation emitted by protons compared to electrons, the LHC requires only 16 accelerating cavities, which are all located at Point 4; in contrast 416 cavities were required for the operation of LEP during its last, and highest energy, phase.

To inject bunches of protons into the LHC for acceleration, a series of pre-accelerators is used. Hydrogen atoms are first stripped of their electrons to produce protons and subsequently accelerated to an energy of 50 MeV by a linear accelerator (Linac2). They are then injected into the Proton-Synchrotron Booster (PSB) followed by the Proton Synchrotron (PS), where they are accelerated to 25 GeV. From there, the Super Proton Synchrotron accelerates them to an energy of 450 GeV before they are injected into the LHC. In 2010 and 2011, the LHC reached an energy of 3.5 TeV per beam and in 2012 this was increased to 4 TeV per beam. An overview of the CERN accelerator complex is shown in figure fig. 2.1.

The minimum distance between bunches in the LHC is 25 ns (7.5 m), but

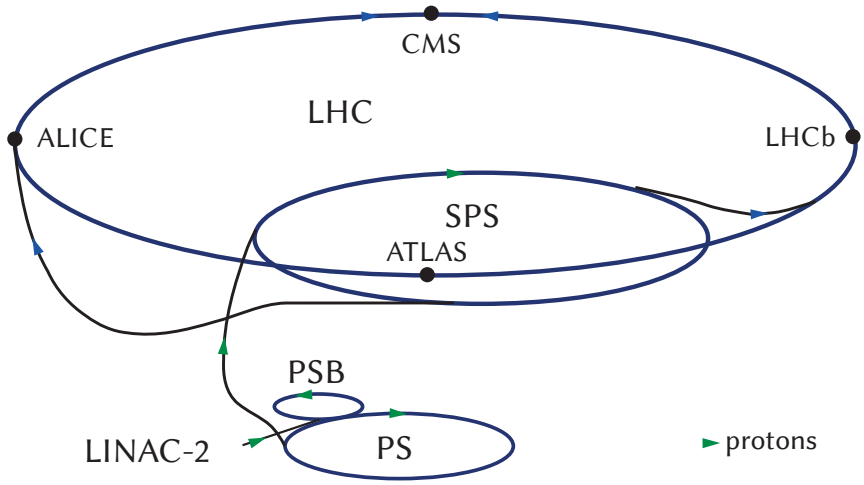


Figure 2.1: Schematic overview of the CERN accelerator complex. The location of the LHC experiments is also shown.

during 2010, 2011 and 2012, a bunch-spacing of 50 ns (15 m) was used, with at most 1331 bunches consisting of approximately 1.15×10^{11} protons each, circulating in both beams. Approximately 1300 pairs of bunches collided at Interaction Point 8 (IP8). Filling of the LHC by the pre-accelerators takes approximately 15 minutes and acceleration of the bunches to their maximum energy takes another 10 minutes [33]. Once the maximum energy is reached, the beams are focused at the interaction points and finally brought into collision.

2.2 The LHCb Detector

When two protons collide in the LHC, the gluons or quarks they consist of interact. The partons that undergo such an interaction each carry a different fraction of the momentum of the proton. The direction of this momentum difference between the interacting partons points along the beamline and as a result, particles that are created in the interaction are boosted along this direction.

Heavy flavour hadrons are produced copiously at the LHC and due to their relatively small mass (several GeV/c^2) with respect to the momentum difference (up to several TeV) of the interacting partons, they are strongly boosted, and thus predominantly produced in the forward direction, as illustrated for b quarks

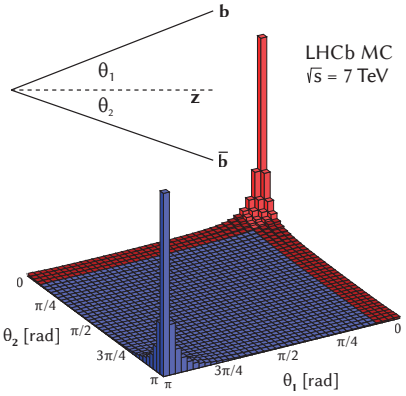


Figure 2.2: Angle with respect to the beamline at which $b\bar{b}$ pairs are produced at $\sqrt{s} = 8$ TeV. The acceptance of the LHCb detector is indicated in red.

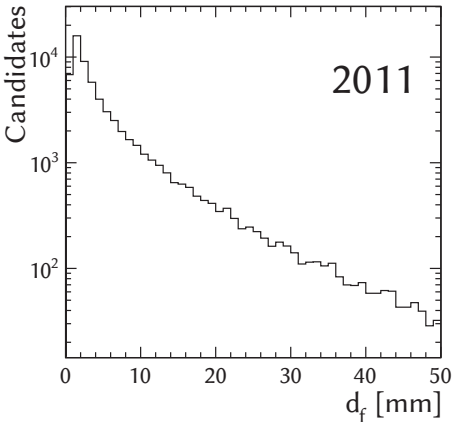


Figure 2.3: Flight distance, d_f, of a sample of $B_s \to J/\psi\,\varphi$ decays recorded in 2011.

by fig. 2.2. The LHCb detector is therefore designed as a single arm forward spectrometer covering the region between 10 and 300 mrad in the bending plane and between 10 and 250 mrad in non-bending plane.

The LHCb physics program focuses on decays of charm and beauty hadrons. One of the properties of such heavy-flavour hadrons, that allows them to be distinguished from background, is their large lifetimes of, for example, 0.41 ps for the D^0 and 1.51 ps for the B_s, respectively [14]. Since such particles are produced with a large boost, the combination with high average lifetimes results in flight distances of at least several millimetres for the majority of the decays, as can be seen in fig. 2.3, for the signal channel of this thesis.

For most of the tracks of charged particles into which heavy flavour hadrons decay, this results in a sizeable distance of closest approach (DOCA), or impact parameter (IP), to the primary interaction vertex. The precise reconstruction of such tracks allows heavy flavour decays to be efficiently distinguished from background tracks and results in small uncertainties on the position of the decay vertices. Since the position of the decay vertex is the largest contribution to the uncertainty on reconstructed decay-times, small uncertainties on the reconstruction of its position additionally lead to high resolution on reconstructed

decay-times. This is studied in chapter 6. The LHCb vertex locator (Velo) detector is positioned around the interaction point and designed to allow reconstruction of charged particle tracks with the required position resolution to resolve fast B_s oscillations. It is described in more detail in section 2.3.

To allow the estimation of the momentum of charged particle tracks, a large dipole magnet is placed approximately five metres downstream of the interaction region, preceded and followed by a set of tracking detectors. The magnet has a bending power of approximately 4.2 Tm, and operates at room temperature. The magnet figures prominently in fig. 2.4, which shows a schematic side view of the LHCb detector. The tracking detectors are described in section 2.4.

To provide identification of charged pions, kaons and protons, two ring imaging Čerenkov (RICH) detectors are installed, one upstream and one downstream of the magnet, adjacent to the tracking stations. Electrons and photons are identified by the silicon pad detector and the preshower and electromagnetic calorimeters, while the muon stations allow identification of traversing muons. The RICH detectors are described in section 2.7, the calorimeters in section 2.8 and the muon detector in section 2.9.

Fast, fixed-latency, hardware based triggers are supplied by the electromagnetic and hadronic calorimeters based on high transverse energy clusters, and by the muon detector based on stand-alone reconstructed track segments; they are described in section 2.10.1. Events accepted by the hardware-based L0 trigger are sent to a large server farm for processing by the software-based high-level trigger (HLT), which is presented in chapter 3.

2.3 The Vertex Locator

To achieve the required impact parameter and vertex position resolution, the Velo is a fine grained silicon strip detector with low material budget to limit multiple scattering. To allow fast pattern recognition, the strips are oriented in the r and ϕ directions, that respectively measure the azimuthal and radial coordinates of charged particle tracks traversing the detector. The detector is divided in two halves, each consisting of 21 modules mounted perpendicular to the beam. The detector halves are contained in a secondary vacuum that is separated from beam vacuum by a 300 µm thick aluminium foil, the so-called RF foil; named such, as it shields the detector from possible radio-frequency (RF) pick-up from the LHC beam. The number of modules is chosen such that tracks that are inside the acceptance of the other LHCb tracking detectors and originate up to 10 cm downstream of the interaction point traverse at least 3 silicon strip

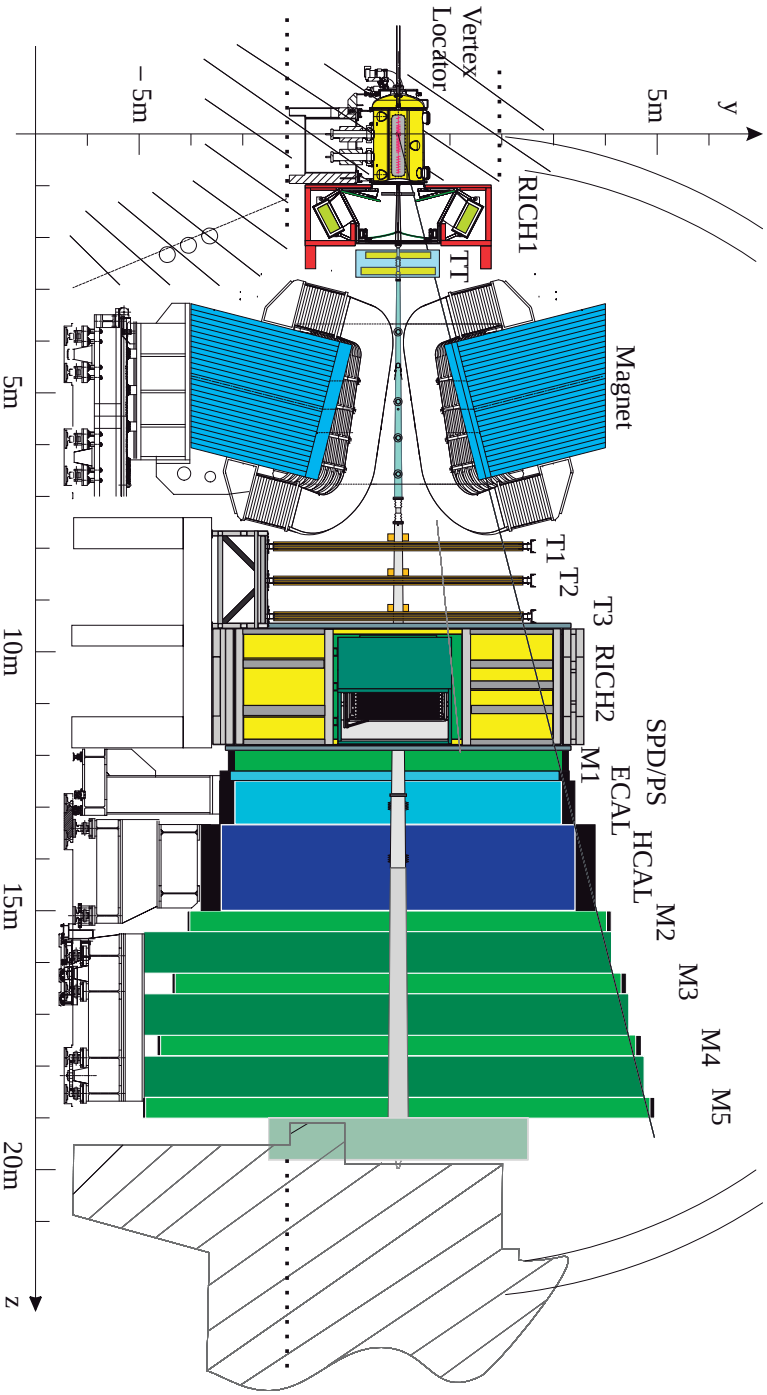


Figure 2.4: Schematic side view of the LHCb detector; the side facing the centre of the LHC is shown.

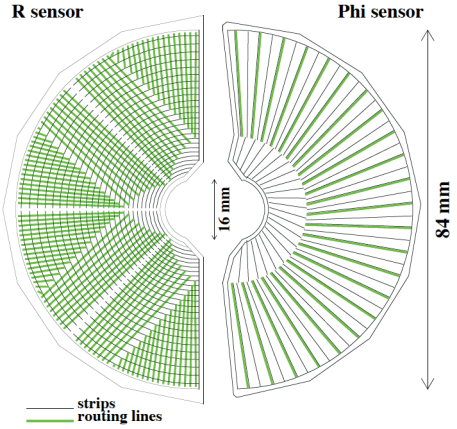


Figure 2.5: A schematic overview of a velo R and Φ sensor.

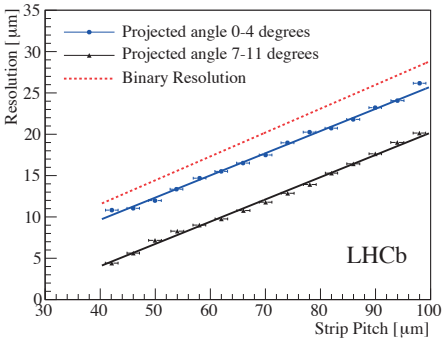


Figure 2.6: Single hit resolution in the Velo as a function of the strip pitch for the R sensors. The resolution is given in two ranges of the projected angle, and the estimated single-hit binary resolution is given for comparison. Figure reproduced from [34].

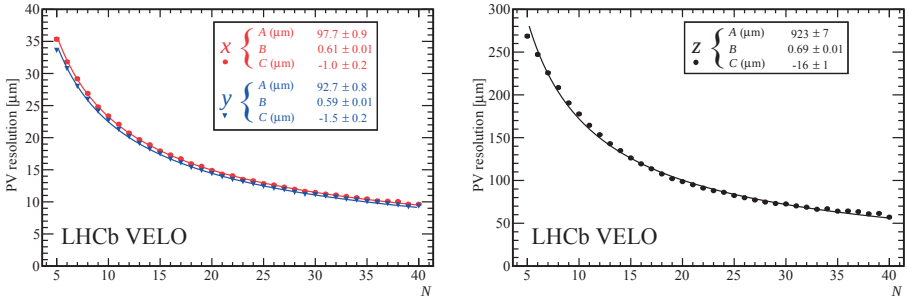


Figure 2.7: Resolution on the reconstructed position of vertices as a function of the number of tracks in the vertex. Figure reproduced from [34].

detector modules.

The strips are located on 300 µm thick silicon sensors, mounted back-to-back on both sides of a carbon substrate to form a detector module. Each sensor consists of 2048 strips. The R-sensor — containing the circular oriented strips — is segmented in 45° quadrants, which have a pitch that linearly increases from 40 µm at the inner edge to 101.6 µm at the outer edge. The Φ-sensor — containing the radially oriented strips — consists of 683 short inner strips and 1365 long outer strips. The inner and outer strips are positioned at a stereo angle between -10° and 10° with respect to the radial direction. Read-out and power infrastructure are connected to the sensors on their outer edge, outside of the detector acceptance. Figure 2.5 shows a schematic overview of a Velo R and a Velo Φ-sensor.

In addition to the measurement of the presence of charge deposits in strips, the Velo also records the amount of charge deposited with an 8 bit ADC. Up to four neighbouring strips containing charge deposits are combined in clusters and the recorded ADC values are used to calculate a weighted average position. This results in a hit resolution that is better than the binary resolution expected from the strip pitch. The number of adjacent strips that a charged particle traverses depends on the so-called projected angle: the angle between the direction of the particle and the normal to the sensor plane. Figure 2.6 shows the hit resolution of the R-sensors, obtained from 2010 data. The best hit resolution is measured to be 4 µm for hits in the region with smallest strip pitch of 40 µm.

The uncertainty on the impact parameter of a charged particle track is dominated by the extrapolation of the track to the location of the primary vertex (PV). The distance over which a tracks must be extrapolated depends on the angle of the track with respect to the beamline and the distance of the first measurement to the beamline. For a given minimum distance of the first hit to the beamline, the uncertainty on the IP is dominated by the resolution on the two closest hits to the PV [35, p. 69]. The reconstructed PV resolution is shown in fig. 2.7 as a function of the number of tracks that constitute a PV, N. The resolution is fitted with a function that parameterises it in terms of N as follows:

$$\sigma_{\mathrm{PV}} = \frac{A}{N^B} + C, \tag{2.1}$$

where A, B and C are constants. The parameter estimates and the corresponding function resulting from the fits to the resolution the x, y and z directions are given in fig. 2.7. A PV consisting of 25 (4) tracks has a resolution of 13 (41) µm in the transverse plane and 71 (340) µm in z.

To minimise the distance over which tracks need to be extrapolated, the Velo is designed to be positioned as close as 8 mm to the beamline. Positioning a

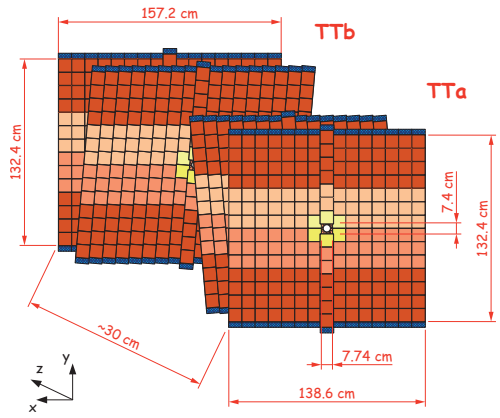


Figure 2.8: Layout of the TT with consecutive layers oriented at 0°, -5°, 5° and 0°.

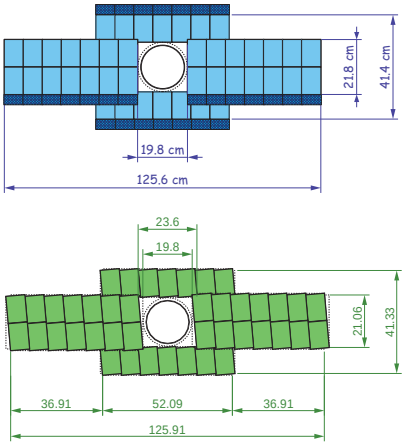


Figure 2.9: Layout of IT layers oriented at 0° (top) and -5° (bottom).

detector at such a small radial distances from the beam incurs several safety risks. These risks are mitigated by designing the detector to be movable, from 30 mm away from the beam during injection down to its operating distance of 8 mm once stable beam conditions are declared. The position of the interaction region can change from fill to fill and the possibility to move the detector allows it to be centred around the interaction region with a precision of a few µm. This is achieved by reconstructing vertices with each halve of the Velo separately and centring the detector around the beam once its position is stable. The halves are then moved together in several steps, interleaved with checks on the stability of their position relative to the interaction region. The closing procedure is fully automated and takes between 5 and 10 minutes depending on the availability of minimum bias events from the HLT.

2.4 Silicon Tracker

The design of the tracking detectors surrounding the beam downstream of the Velo depends on the average density of charged particle tracks, which is approximately uniform as a function of rapidity. It strongly increases in the forward direction as the beamline is approached, and to ensure efficient pattern recognition with a low ghost rate, higher granularity detectors are required in this

region. Silicon strip detector are, therefore, installed close to the beamline. Upstream of the magnet this is the Tracker Turicensis, or TT, and downstream of the magnet it is the inner tracker, or IT.

The TT is a silicon strip detector consisting of four planes of 500 μm thick silicon strips with a pitch of 183 μm, covering a total area of 8 m^2. The strips on these four consecutive planes are oriented at 0°, -5°, 5° and 0° with respect to the vertical direction to aid pattern recognition of charged particle tracks; fig. 2.8 shows a schematic overview of the layout of the TT. The layers oriented at 0° are called vertical layers, while those oriented at -5° and 5° are called stereo layers. The fraction of correctly functioning channels is high, but varied somewhat during LHC Run I, related to issues with the readout. The construction of the TT allowed repairs to be made during short, week-long, technical stops, and the luminosity weighted average percentage of working channels is calculated to be 99.7%. The occupancy of the TT was found to vary from 0.2% at the outside of the detectors to 1.9% for the strips closest to the beamline. Using a clear sample of $J/\psi \to \mu^+\mu^-$ decays, the hit efficiency was found to be 99.7% with a single hit resolution of 52.6 μm [36].

The IT consists of three stations of four layers each, and covers a cross-shaped area with a total surface of approximately 4 m^2 in the middle of the downstream tracking stations around the beamline. The IT strips have a pitch of 198 μm and the sensors located above and below the beamline are 320 μm thick, while those on its left and right side are 420 μm thick, due to the usage of two-module ladders. Figure 2.9 shows an overview of the arrangement of the IT sensors. The construction and position of the IT only allowed repairs to be made during the long LHC shutdowns at the end of the year and the luminosity weighted average percentage of working channels is lower than for the TT at 98.6%. The occupancy of the IT is the same as the TT at 0.2% for the outside modules and 1.9% for those closest to the beamline. From the same sample of $J/\psi \to \mu^+\mu^-$ decays used for the determination of the hit efficiency in the TT, the IT hit efficiency is found to be 99.8%, with a hit resolution of 50.3 μm [36].

2.5 Outer Tracker

Since the cross-sectional area of the LHCb acceptance downstream of the magnet is too large for an all silicon tracking detector to be affordable, the inner tracker is surrounded by a large area straw-tube detector: the outer tracker, or OT. Approximately 70% of the charged particle tracks that are produced inside the LHCb acceptance pass through the OT. The OT consists of four layers of modules per

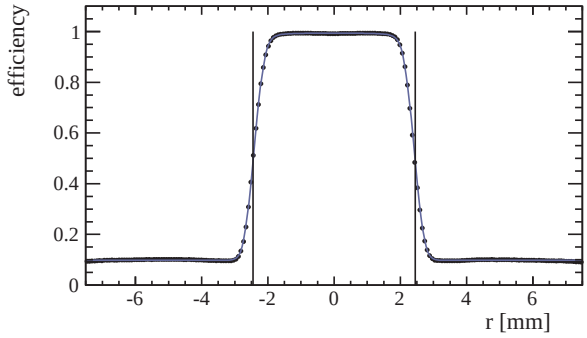


Figure 2.10: Hit efficiency profile for a single outer tracker straw as a function of the distance to wire at the centre of the straw. The vertical likes indicate the edges of a cell.

tracking station, arranged at the same 0°, -5°, 5° and 0° degree angles as the TT and IT. Each module contains two staggered layers of circular straws with an inner diameter of 4.9 mm. The straws are filled with a mixture of Argon (70%), CO_2 (28.5%) and O_2 (1.5%) to ensure a drift time below 50 ns. The straws have a single hit resolution of 205 μm [36].

Figure 2.10 shows an example of a hit efficiency profile for a single straw in the OT. The lower efficiency closer to the edge of the straw is the result of two effects: the lower probability for an ionisation to occur in the gas, due to a charged particle traversing a smaller volume of gas near the edge; and the positioning of a fraction of the hits outside of the straw due to the uncertainty on the extrapolated position of charged particle tracks. The horizontal tails in the neighbouring chambers are caused by other tracks; in other words, they are related to the occupancy. The average hit efficiency for the volume within 1.25 mm from the wire is found to be 99.2%. The occupancy of the OT straws on the outside of the detector was found to be 4%, while for the straws closest to the beamline it was as high as 17%. This higher occupancy than the TT and IT is a result of the larger size of the straws and of the fact that a 75 ns wide read-out windows is employed to ensure all hits are collected over the full drift-time spectrum.

2.6 Tracking

Charged particle tracks reconstructed in the LHCb tracking detectors are classified into five categories, depending on which tracking detectors they traverse:

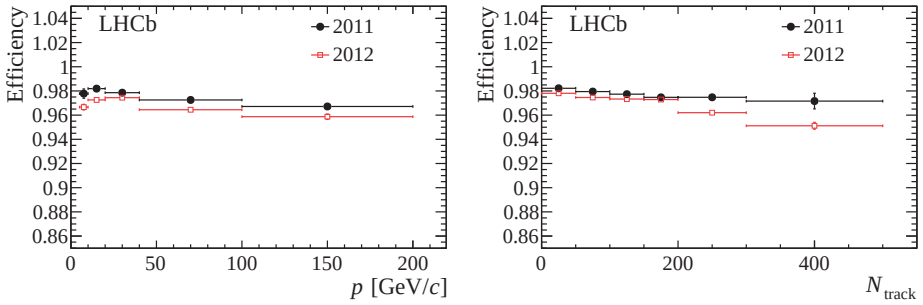
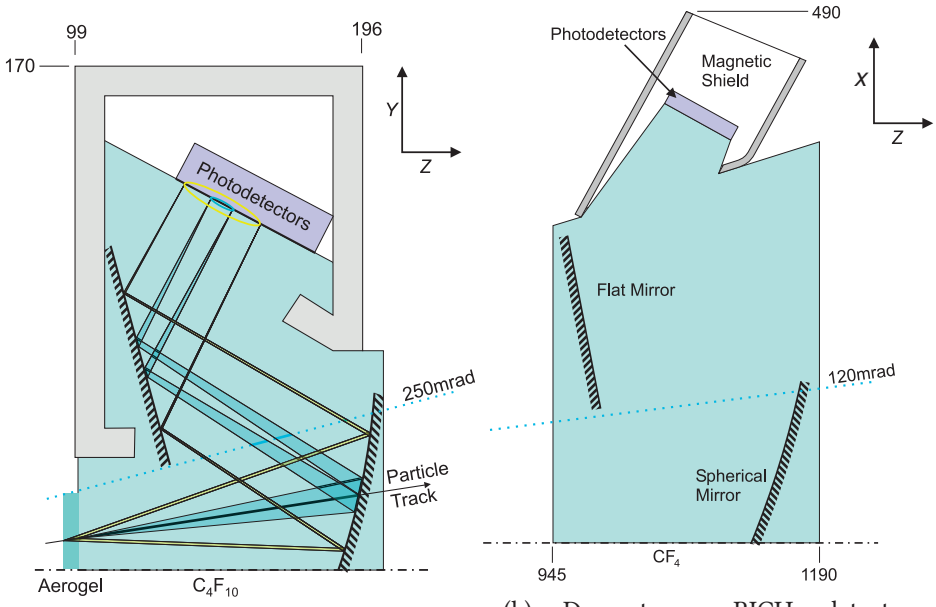


Figure 2.11: Long track reconstruction efficiency as a function of estimated track momentum, p, and number of tracks in the event, N.

- tracks that traverse only the Velo are called Velo tracks,
- tracks that traverse the Velo and the IT or the OT are called long tracks,
- tracks that traverse the Velo and the TT, but neither the IT nor the OT are called upstream tracks,
- tracks that are produced outside the Velo but traverse the TT and the IT or the OT are called downstream tracks,
- and tracks that solely traverse either IT or the OT are called T tracks.

The majority of physics analyses use long tracks exclusively, while downstream are used for analyses involving long-lived intermediate resonances such at K^0_S or Λ^0. Velo tracks are only used for primary vertex reconstruction, upstream tracks can be slow pions and T tracks often secondaries.

Tracks are reconstructed by a number of pattern recognition algorithms and subsequently fitted using a Kalman-filter [37] based approach. The first pattern recognition step is the Velo track finding, which is described in more detail in the context of the HLT in section 3.2.1. An algorithm called the forward tracking subsequently attempts to extend Velo tracks to long tracks by searching for hits in the downstream tracking stations; this algorithm is described in more detail in section 3.2.4. Independently, an algorithm, referred to as track seeding, creates track segments exclusively using hits in the tracking stations. These track segments are matched to Velo tracks to create additional long tracks. Long tracks created by the forward and the combination of seeding and matching are compared and duplicates are removed. Unused tracking station segments are then extended to the TT to create downstream tracks, provided sufficient additional TT hits are found. Unused Velo tracks are analogously extended to the TT to create upstream tracks. As a result of the high single hit efficiencies of all track-



(a) Upstream RICH detector (RICH1)

(b) Downstream RICH detector (RICH2)

Figure 2.12: Schematic overview of the LHCb RICH detectors installed upstream and downstream of the magnet. The path of the light shown in the figure is that of photons emitted in the plane formed by the direction of the momentum of the radiating particle and the vertical axis; reproduced from [38, p. 41].

ing detectors, an average long track finding efficiency of above 96% is realised, as is shown in fig. 2.11. Finally, the Kalman track fit is applied to all reconstructed tracks to create the final set of tracks; this is described in section 3.2.6. For Velo tracks, the fit is a simple straight line fit, as they don't traverse any magnetic field.

2.7 The RICH Detectors

Charged particles travelling through a medium at velocities larger than the speed of light in that medium emit so-called Čerenkov radiation in the visible spectrum [39]. This light is emitted at a an angle, θ, who's cosine is inversely proportional

to the particle's velocity, β, according to:

$$\cos\theta = \frac{1}{n\beta}, \tag{2.2}$$

where n is the refractive index of the material. A measurement of the Čerenkov angle thus yields the velocity of a charged particle which, when combined with a measurement of its momentum, can be turned into a determination of its rest mass, m_0. To allow identification of particles with momenta in the range of 2 $-$ 100 GeV/c, two ring imaging Čerenkov (RICH) detectors are installed in the LHCb experiment, containing radiator volumes with different refractive indices.

The first detector, called RICH1, is installed upstream of the magnet directly after the Velo, and provides charged particle identification for particles with momenta in the range of 2 $-$ 50 GeV/c. The second detector, called RICH2, is installed downstream of the magnet between the tracking stations and the calorimeters. It provides identification of charged particles with momenta up to 100 GeV/c. Both detectors consist of two halves that contain: a radiator volume, a set mirrors and an array of photo detectors arranged in the focal plane of the system of mirrors. The photo detectors of RICH1 are placed above and below the beamline, while they are placed on its left and right side in RICH2. The RICH1 radiator volume consists a 5 cm thick volume of silica aerogel, $n = 1.03$, installed at the entrance of a 85 cm long C_4F_{10} volume with $n = 1.0014$. The RICH2 radiator is a 167 cm long volume filled with CF_4, $n = 1.0005$. The first, spherical, mirror focuses the Čerenkov light emitted by a charged particle into a ring in its focal plane. The second, flat, mirror reflects the focused light onto the photo detectors.

To identify charged particles, the size and position of the ring of emitted Čerenkov light on the photo detectors is calculated, using the measured momentum, for the hypotheses that a particle is a pion, kaon or proton. The likelihood of each of these hypotheses given the photon hit data is then calculated and stored. Selection requirements can subsequently use the difference in the logarithm of these likelihoods (-$\Delta\log\mathcal{L}$) to select particles that are, for example, more likely to be kaons than pions or protons. The type and refractive indices of the radiating media in RICH1 and RICH2 were chosen to optimise the separation of the measured rings of Čerenkov light emitted by different types of particles with momenta as given above, thereby maximising the -$\Delta\log\mathcal{L}$ between the pion, kaon and proton hypotheses.

Figure 2.13 shows the kaon identification and pion mis-identification probabilities provided by the the RICH detectors for charged particles with momenta up

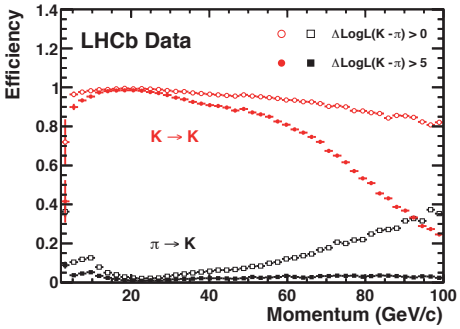


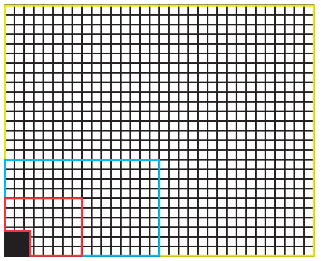
Figure 2.13: Combined RICH1 and RICH2 kaon identification and pion misidentification probabilities as a function of momentum for two different requirements on the difference between the logarithms of the likelihood for being either a kaon or a pion; figures reproduced from [36].

to 100 GeV/c, obtained from a sample of $D^{*+} \to D^0(\to K^-\pi^+)\pi^+)$ decays, where the charge correlation between the second π and the two remaining hadrons is used to identify them as a kaon and pion, respectively. To illustrate the effect of different selection requirements on the -$\Delta\log\mathcal{L}$ between being a kaon and a pion, the probabilities are shown for two -$\Delta\log\mathcal{L}$ values. For the requirement of -$\Delta\log\mathcal{L}(K—\pi) > 0$, and averaging over the momentum range 2–100 GeV/c, the average probability of correctly identifying a kaon is $\sim$ 0.95 with an average pion mis-identification probability of $\sim$ 0.1. For the stronger requirement of -$\Delta\log\mathcal{L}(K—\pi) > 5$, which reduces the average mis-identification probability to 0.03, the average identification probability is reduced by approximately 10% to 0.86.

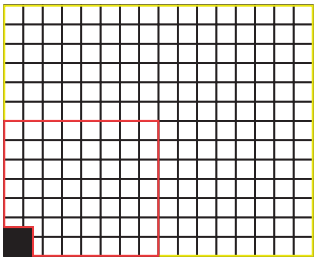
2.8 The Calorimeters

The LHCb calorimeter system consists of a scintillating pad detector (SPD), a preshower detector (PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). They are installed downstream of the second RICH detector and provide identification of electrons and photons, and fast triggers based on high transverse energy clusters.

To facilitate association of energy deposits by the same particle in different calorimeters, all calorimeters are designed projectively toward the interaction point. Since the particle density varies over two orders of magnitude from the



(a) Segmentation of the SPD, PS and ECAL; the cell sizes for the SPD and PS are 1.5% smaller than those given here for ECAL.



(b) Segmentation of the HCAL detector.

Figure 2.14: Segmentation of a single quadrant of the LHCb electromagnetic and hadronic calorimeters.

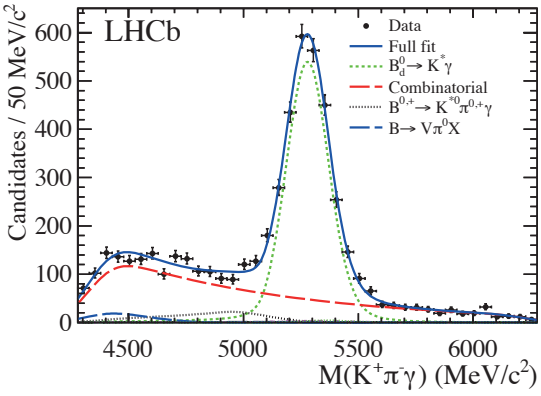


Figure 2.15: Distribution of $M_{K\pi\gamma}$ for a sample of decays of $B^0 \to K^{*0}\gamma$ selected from 1 fb^1 of data collected by LHCb in 2011; reproduced from [40].

inside to the outside, calorimeter cell sizes are increasingly smaller toward the inside. Figure 2.14a shows the segmentation of the SPD, PS and ECAL. The blocks shown in the inner, middle and outer regions contain nine cells, four cells and a single cell, respectively.

Both ECAL and HCAL are sampling calorimeters based on so-called shashlik technology [41, 42]. The ECAL uses lead tiles as the absorber material, while the HCAL uses steel. A particular feature of the design of the HCAL is that the scintillating tiles are oriented parallel to the beamline, while they are perpendicular to it in the ECAL.

The ECAL is 25 radiation lengths (X_0) thick to fully contain electromagnetic showers initiated by photons and electrons. It is mainly used for reconstruction and identification of photons and electrons originating from heavy flavour decays. Additionally, it provides the highest transverse energy clusters as input to the hardware-based L0 trigger. Figure 2.15 shows the invariant mass resolution for a selected sample of decays of $B^0 \to K^{*0}\gamma$. The resolution on the mass of the B^0 is 93 MeV/c^2 in this case, which is dominated by the resolution on the reconstructed photon energy.

The largest background for high transverse energy electron and photon triggers consists of charged hadrons. To be able to remove this background, the preshower detector (PS) is installed. It consists of a layer of scintillating pads behind a layer of lead absorber with a thickness of $2X_0$. Since this thickness of lead represents only approximately 6% of an hadronic interaction length, electrons and photons are much more likely to induce a shower in this material than charged hadrons. Electrons and photons are thus identified by the hardware trigger as an energy deposit above threshold in both the PS and ECAL detectors.

To distinguish between electrons and photons, an additional layer of scintillating pads (SPD) is installed in front of the lead absorber of the PS. Electrons that traverse this layer are likely to produce a signal, in contrast to photons. Additionally, the SPD provides a measure of the event multiplicity to the hardware trigger.

Due to space limitations, and to less stringent requirements on the resolution of reconstructed clusters than for ECAL, the HCAL is only 5.6 interaction lengths (λ_i) thick. Its segmentation is coarser than the ECAL and shown in fig. 2.14b. Given its thickness, most hadronic showers will not be fully contained and result in significant amounts of punch-through to the muon stations. This is however not considered problematic, as the main purpose of the HCAL is to provide triggers on high transverse energy clusters.

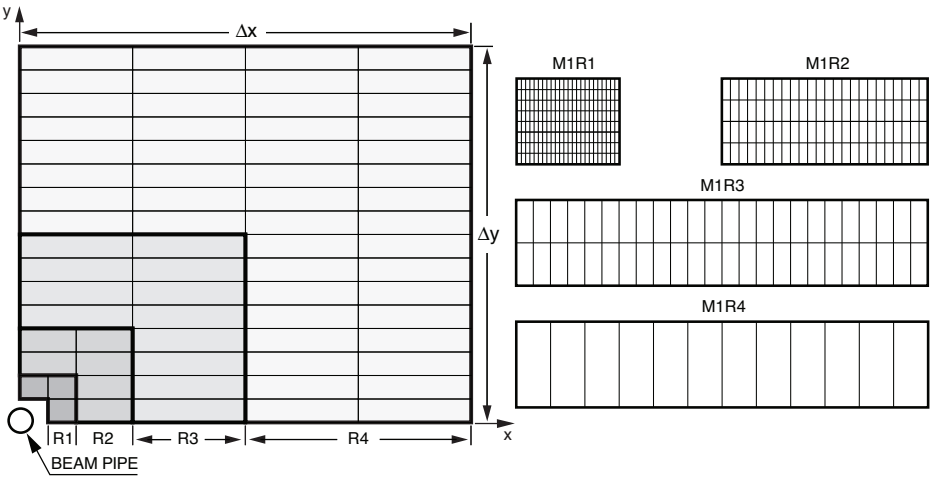


Figure 2.16: Left: schematic overview of the arrangement of multi-wire proportional chambers in a muon station. Right: overview of the segmentation into readout pads of the chambers in muon station M1. In the corresponding regions of stations M2 and M3 (M4 and M5) contain double (half) the number of columns of pads, while the number of rows is the same [43].

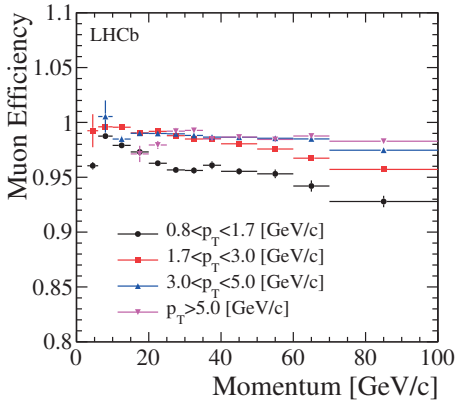


Figure 2.17: Probability of correctly identifying a reconstructed charged particle track as a muon for difference transverse momenta, as a function of track momentum; reproduced from [36].

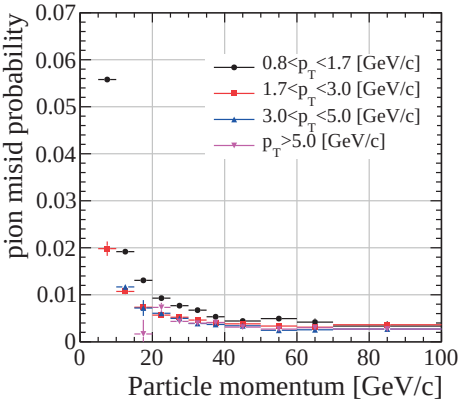


Figure 2.18: Probability of misidentifying a reconstructed charged pion track as a muon for different transverse momenta, as a function of track momentum; reproduced from [36].

2.9 The Muon System

The final LHCb subdetector in the spectrometer is the muon detector, which is installed downstream of the calorimeters. It is equipped with a large number of multi-wire proportional chambers (MWPCs). The exception is the inner region of the first station, M1, which is installed in front of the calorimeters, and is equipped with Triple-GEM detectors. The stations M2–M5 are interspersed with layers of iron shielding to absorb punch-through. The muon detector provides offline identification of muons, and fast muon triggers based on reconstructed track segments.

Similar to the calorimeters, the muon stations are projectively segmented with respect to the interaction point to facilitate fast reconstruction of track segments [43]. Figure 2.16 shows the arrangement of MWPCs on a quarter of a muon station. The projectivity is particularly useful for the hardware-based trigger, where track segments with hits in all five muon stations are required.

The muon identification algorithm starts from reconstructed tracks and calculates a likelihood of these being a muon by searching for associated hits in the muon stations; the algorithm is described in more detail in section 3.2.5. The likelihood method is calibrated using samples of $J/\psi \to \mu^+\mu^-$ and $\Lambda \to p\pi^-$ for the muon and not-muon likelihoods, respectively. Figure 2.17 shows the correct identification probability and the probability that a pion is mis-identified as a muon obtained from the likelihood method. The probability of mis-identifying a pion includes decays-in-flight.

2.10 The LHCb Trigger

At an instantaneous luminosity of 4×10^{32} cm^2/s with an average number of 2.5 proton-proton interactions per beam crossing provided by the LHC [44], the rate of beam crossings that result in at least one visible interactions in the LHCb detector is approximately 12 MHz. The purpose of the LHCb Trigger is to select from these 12 MHz of crossings, those that are suitable for LHCb physics analysis with high efficiency and minimal bias, while rejecting as much background as possible.

The trigger is structured as a two-stage system, which is schematically shown in fig. 2.19. Since the full detector can be read out at a maximum rate of 1 MHz, the first stage (L0) is implemented in dedicated hardware to provide an initial reduction of the event rate down to this limit; it is based on data from the calorimeters and the muon stations and has a fixed latency of 4 μs. It is described below.

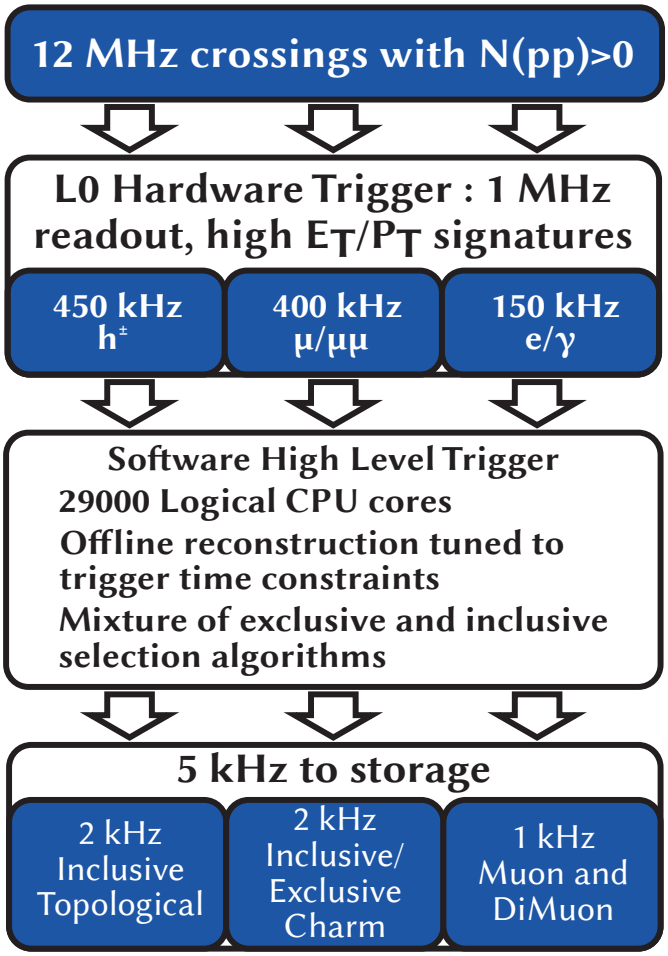


Figure 2.19: Schematic overview of the LHCb trigger in 2012.

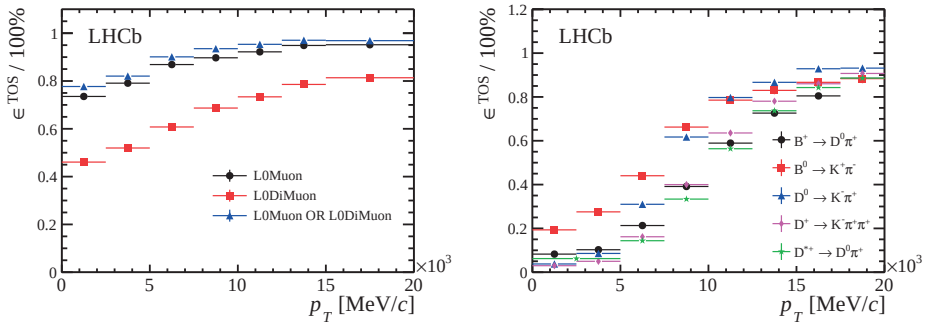


Figure 2.20: Left: L0Muon TOS efficiency for selected $B^+ \to J/\psi K^+$ candidates. Right: L0Hadron TOS efficiency for a number of beauty and charm decays.

The second stage is called the High Level Trigger (HLT) and consists of a software application running on a PC-server farm; it is introduced in section 2.10.2 and described in detail in the following chapter.

2.10.1 The Hardware Trigger

The L0 trigger consists of a number of dedicated hardware boards that process raw data from the muon stations, the calorimeters and the so-called pile-up detector. The information provided by each of these systems is combined with the event multiplicity measured by the SPD detector in a programmable decision unit (L0DU) to obtain the final decision [45].

The calorimeter based L0 trigger discriminates signal events from background based on the transverse energy of clusters in the ECAL and HCAL [46, Chapter 2]. The main trigger for beauty and charm decays to fully hadronic final states is provided by placing a threshold on the minimum transverse energy of the highest energy HCAL cluster. The right side of fig. 2.20 shows the efficiency of the L0 hadron trigger. The efficiency displayed is the so-called TOS efficiency, which is defined in section 3.1. The highest average efficiency of 40% is achieved for two-body beauty decays, while multi-body charm decays are triggered with the lowest average efficiency of 20%.

L0 triggers for decays of beauty hadrons to final states including photons and electrons are provided by a combination of the SPD, PS and ECAL. The L0 triggers for electrons and photons are based on the largest ECAL cluster being above threshold with or without a hit in the SPD pad in front of the ECAL cluster.

The energy deposited in the PS in front of the ECAL cluster must additionally be above threshold to remove background from charged hadrons.

The muon based L0 triggers reconstruct track segments in muon stations M1 to M5 using a lookup table based algorithm [46, Chapter 3]. Due to the availability of a hit in M1, the momentum is estimated with an uncertainty of approximately 25% [47]. The single muon L0 trigger accepts events if at least one candidate is found with a p_{T} exceeding the threshold. The dimuon trigger accepts events where the product of the p_{T} of the two highest p_{T} candidates exceeds the threshold. The threshold is tuned such that the p_{T} of both muons may be below the threshold of the single muon trigger. The left side of fig. 2.20 shows the TOS efficiency of the L0 muon based triggers. For selected $\mathrm{B}^+ \to \mathrm{J}/\psi\mathrm{K}^+$ decays, the average efficiency of the combination of the single muon and dimuon triggers is 89%.

The pile-up detector consists of an additional four Velo sensors in the backward direction. These sensors are only instrumented with R sensors and are not used for Velo track finding. Instead, their data is used as input to a vertexing algorithm implemented in hardware, which is able to distinguish between zero, one, two, or more than two vertices in the event [48]. This information was initially intended to allow vetoing of events with more than one primary interaction, but was finally only used as part of the online measurement of instantaneous luminosity.

The L0DU combines the data from the individual triggers described above to provides the hardware trigger decision. The values of the different thresholds and how individual triggers are combined is programmable and a specific configuration is tagged with a unique identifier called the L0 trigger configuration key, or L0TCK.

2.10.2 High Level Trigger

The data accepted by the hardware trigger is passed to the second stage of the trigger, the high level trigger (HLT). The HLT is implemented as a software application that runs on the LHCb online PC-server farm, and reduces the rate of events saved to storage to a manageable level, as defined by the computing model; originally in [49, Sec. 4.4.2], with subsequent updates in [50, 51, Sec. 3]. In 2011 the output rate of the HLT was increased from 2 kHz to 3 kHz, while in 2012 it was set to 5 kHz. The HLT and its performance are described in detail in the following chapter.

Chapter 3

The LHCb High Level Trigger

The LHCb detector can be read out at a maximum rate of 1 MHz and at most several kHz of events can be stored on tape. The purpose of the LHCb High Level Trigger (HLT) is to select with high efficiency, events that are interesting for physics analysis at a maximum output rate of several kHz; 2 kHz in 2011 and 5 kHz in 2012.

The large cross-sections for $c\bar{c}$ and $b\bar{b}$ production in the acceptance of the LHCb detector at the LHC of 1419 ± 133 μb [52] and $1.14 \pm 0.01 \pm 0.16$ μb [53], respectively, combined with the constant instantaneous luminosity in the LHCb interaction point of 4×10^{32} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ result in an event rate of approximately 600 kHz of $c\bar{c}$-pair and 20 kHz of $b\bar{b}$-pair production in the LHCb acceptance. The HLT must therefore not only select $b\bar{b}$ and $c\bar{c}$ events out of the 1 MHz L0 rate, but in addition selectively filter specific signal decays out of the large rate of $b\bar{b}$ and $c\bar{c}$ interactions.

The HLT application runs on the online CPU farm, which consisted of 20000 (25000) CPU cores in 2011 (2012). The number of CPU cores available in the online farm, together with the 1 MHz rate of events coming from the L0 trigger determine the time budget for the HLT. For 2011 approximately 25 ms were available on average per event, while in 2012 the budget increased to 33 ms per event due to the implementation of the so-called deferred HLT.

The deferral process increases the effective size of the HLT farm by using the otherwise idle time between fills to process events. In the deferred HLT, the local hard disks of the HLT nodes are used as temporary storage for events accepted by the L0, that cannot be immediately processed by the HLT. These events are processed once the fill ends. If the time between the end of a fill and the start of the next fill is not sufficient to process all deferred events, less disk space is

available for deferral during the next fill. To avoid running out of disk space for deferral, the percentage of events to be deferred in 2012 was chosen to be approximately 25%, based on the expected average duration of an LHC fill, the expected time between fills and the available disk space — approximately 1 TiB per HLT node in 2012.

To optimise the time consumption of the HLT process, it is split into two stages, HLT1 and HLT2, which are discussed in sections 3.2 and 3.3. The first stage, HLT1, contains six optimised and inclusive selections. These selections are based either on single detached tracks, so-called due to their high impact parameter to all interaction vertices, or on pairs of identified muons. To minimise the number of performed reconstruction steps, selection requirements are imposed as early as possible. For example, minimal (transverse) momentum is required of reconstructed tracks as soon as the pattern recognition, that extends Velo track segments to the tracking stations, has completed, before performing the CPU intensive track fit.

The purpose of the HLT1 selections is to reject background as fast as possible and to lower the rate of accepted events such that reconstruction of all tracks traversing the magnetic field is feasible at the start of the second HLT stage, HLT2. The reconstruction of all tracks is followed by inclusive selections based on dimuons or combinations of multiple charged particles, complemented by exclusive selections, for a total of approximately 200 selections. These selections are discussed in sections sections 3.3.2 and 3.3.3. A number of these selections perform additional reconstruction steps and particle identification steps, such as reconstruction of neutral particles and identification of muons, photons and electron. The additional reconstruction steps are described in more detail in section 3.3.1.

The CPU time budget available to HLT1 and HLT2 is summarised in table 3.1. The rate reduction of an approximate factor 12.5, to 80 kHz, by HLT1 results in a larger effective time budget per event available for additional reconstruction in HLT2. The reduction factor is a result of a balance between spending additional CPU budget in HLT1 to improve the efficiency of the HLT1 selections, and the consequent loss of efficiency in HLT2 due to its reduced CPU budget. The output rate of HLT1 strongly affects the CPU budget available in HLT2 as well; it is reduced as much as possible without incurring unacceptable losses in efficiency.

All of the HLT selections rely on charged particle tracks, so their efficient reconstruction is of fundamental importance for the performance of the HLT. This is reflected in the fact that approximately 60% of the available online CPU budget is used for track reconstruction algorithms. To keep biases in physics analyses

	combined	HLT1	HLT2
2011	25 [ms]	15 [ms]	125 [ms]
2012	33 [ms]	17 [ms]	200 [ms]

Table 3.1: Time budget of HLT1 and HLT2 in 2011 and 2012.

due to differences between the online and offline reconstruction of tracks to a minimum, the reconstruction algorithms used online and offline are the same, but they have slightly different configurations to meet the CPU requirements of the online environment.

Two different approaches are taken to achieve this: "changes in the configuration" and "on-demand reconstruction". The most significant change in the configuration of the track reconstruction algorithms is that the minimum transverse momentum at which charged tracks are searched for online is set higher than offline. This decreases the size of the search windows in the pattern recognition algorithms, which leads to fewer hits in regions of interest and reduces combinations of hits to be considered leading to significant gains in speed. Furthermore, a simplified detector geometry is used by the track fitting algorithm required to estimate the amount of detector material traversed by a track to calculate the amount of multiple scattering.

3.1 Measuring HLT Performance

The two most important performance metrics of the HLT are the average time required to process an event and the efficiency of the HLT's various reconstruction algorithms and selections.

The average time required to process an event is measured for new configurations by running the HLT application on a small number of online farm PCs and injecting previously recorded events that were accepted by the L0 trigger, but unbiased with respect to the HLT. Once the average time per event is known the rate of events that the full HLT farm can handle, without deferring any events, is obtained from:

$$R_{\mathrm{HLT}} = \frac{N_{\mathrm{HLT}}}{t_{\mathrm{ev}}}, \tag{3.1}$$

where N_{HLT} is the number of clones of the HLT application running on the HLT farm and t_{ev} is the average required to process an event. Including a fraction f_{defer} of events is to be deferred, R_{HLT} must be larger than $(1 - f_{\mathrm{defer}})$ MHz to be able to cope with the rate of events accepted by the L0 trigger during data taking.

The efficiency of the reconstruction steps and selections is evaluated using the "triggered on signal, versus triggered independent of signal" (TISTOS) method. This method utilises the fact that the hits in the detector, used to create a reconstructed object — a track or calorimeter cluster — uniquely define that object. This is also true for single particles, (e.g. a pion), as they are created from tracks or calorimeter objects and for composite particles, (e.g. a B_s) as they, in turn are created from multiple single particles. Lists of hits of online and offline selected objects can be compared to test if these objects are based on identical information.

To be able to do this, the online reconstruction records the set of detector hits for each track, calorimeter cluster, basic particle or composite particle that is accepted by each trigger selection. For HLT1 these are typically single track or two-track combinations, while most HLT2 selections create composite particles. For two-track combinations in HLT1 or composite particles in HLT2, the hits are saved in a tree-like structure classifying which hits belong to which basic particle and which basic particles make-up which composite particles.

When comparing offline reconstructed objects to online reconstructed objects, the tree-like structure of detector hits for the offline objects is compared to each of the structures stored for a trigger selection of interest. For each comparison, all offline branches are considered, and if more than 70% of the hits present on an offline branch is also present on an online branch, the offline branch is marked as triggered on signal (TOS), while if less than 1% of the hits on the offline branch is present on the online branch, the offline branch is marked as triggered independent of signal (TIS). Offline branches which have more than 1% but less than 70% overlap with an online branch are classified as neither TIS nor TOS. If all offline branches are marked as TOS or TIS, with respect to one of the online structures, the offline object is classified accordingly, with respect to the corresponding trigger selection.

If an offline object is classified as TOS with respect to a trigger selection, the presence of the offline reconstructed object is considered to have been sufficient for the event to have been accepted by the trigger selection, i.e. the offline particle "triggered" the event. In case of a TIS classification, the trigger selection is considered to have accepted the event independent of the presence of the offline object under consideration, i.e. the event was selected without the presence of this specific offline object.

The TIS and TOS classifications of an offline reconstructed particle are not necessarily mutually exclusive. If the trigger reconstructed multiple objects in an event and the presence of more than one of these would cause a trigger selection

to accept the event, an offline particle may be classified as TOS with respect to one online object and TIS with respect to one of the others.

Using the TIS and TOS classifications, an efficiency is defined as follows:

$$\varepsilon^{\text{TOS}} = \frac{\varepsilon^{\text{TOS}\wedge\text{TIS}}}{\varepsilon^{\text{TIS}}}. \tag{3.2}$$

To avoid biases, care should be taken to ensure that the phase-space covered by the trigger selection(s) used for the TOS sample is fully covered by the phase-space of the selection(s) used for the TIS sample.

Since the number of offline heavy flavour candidates classified as TIS for a particular trigger selection is generally rather small, a number of trigger selections may be combined to provide a larger TIS sample. The sample of online objects with which an offline particle is compared for a TIS classification is enlarged, which increases the probability of being classified as TIS. When using a larger sample of trigger selections for the TIS classification, care should be taken to avoid biases being reintroduced by the selections of (one of) these additional triggers.

To avoid such biases but still be able to determine the efficiency of at least part of the selection requirements of a trigger selection, an alternative TOS trigger selection can be introduced that does not contain a specific set of selection requirements and replaces these with unbiased alternatives such as a prescales.

3.2 HLT1

To minimise the CPU time spent on the reconstruction of charged particle tracks, the HLT1 consists of a few optimised inclusive selections that interleave selection requirements and reconstruction steps as much as possible. The selection can be divided into four categories: selections which search for single, high-quality, displaced tracks; dimuon selections; high-energy photon and electron selections and so-called technical selections for luminosity and monitoring purposes. In this thesis, the focus is on dimuon selections, and the reconstruction steps available in HLT1 are, therefore, described in the following sections in the order in which they are used in the dimuon selections.

The high efficiency and background rejection, coupled with the relative simplicity of the LHCb muon identification, allows muons of interest to be identified at a very early stage in HLT1. The HLT muon identification algorithms are discussed in sections 3.2.2 and 3.2.5. This enables a high quality reconstruction of

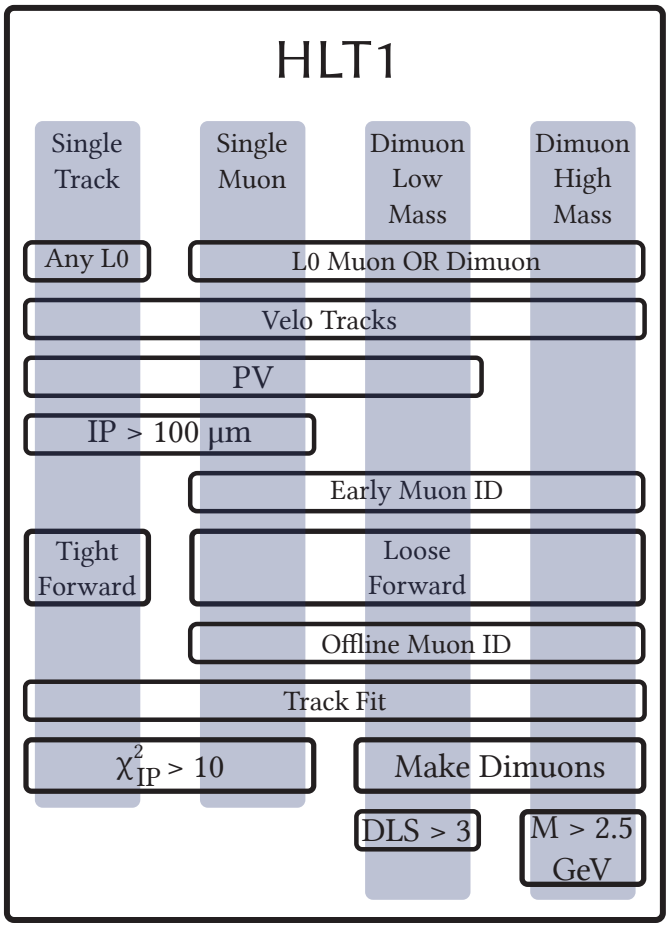


Figure 3.1: Schematic overview of reconstruction steps and important selection requirements applied by HLT1 single track and dimuon selections.

muon tracks, without stringent requirements on momentum and transverse momentum. For dimuon candidates with a mass above 2.7 GeV/c^2 requirements on the displacement of the tracks with respect to the reconstructed primary vertices are not required. This allows a selection, that does not result in a decay-time dependent acceptance (see chapter 5). For dimuon candidates with a mass between 1 and 2.7 GeV/c^2 a displacement requirement is needed to reduce the rate to an acceptable level. The HLT1 inclusive dimuon selections are described in section 3.2.7.

The rate of charm and beauty decays to fully hadronic final states is much higher than the rate of decays to final states including muons. To be able to efficiently select fully hadronic decays in HLT1, the single track selections search for single tracks with low χ^2/DoF, high momentum and high transverse momentum, that are not associated to any of the reconstructed primary vertices. They are described in section 3.2.8.

At an input rate of 1 MHz of events accepted by the L0 trigger, HLT1 accepts events at a rate of approximately 80 kHz. This consists of 60 kHz of single track triggers, 4.3 kHz of dimuon triggers and 16 kHz of other triggers, with some overlap.

The following sections describe the reconstruction steps used as part of the dimuon and single track HLT1 selections, in the order in which they are used by the dimuon selection for candidates with a mass in the range between 1 and 2.7 GeV/c^2. Figure 3.1 shows a schematic overview summarising the HLT1 reconstruction steps used as part of the HLT1 selections, and the most important selection requirements. The HLT1 selections are implemented in the LHCb software by so-called streamers, which are described in more detail in appendix A.

3.2.1 Velo Track Finding

After the 2010 run, a new Velo track finding algorithm was implemented which offered significantly decreased processing times [54]. This allowed the Velo track reconstruction and primary vertex search to be performed for all events as the first step in HLT1. The previous implementation of the Velo track finding algorithms was a three-step procedure where the first step exploited the R-Φ design of the Velo sensors to perform a projection of the hits into *rz* space, followed by a 2D search for tracks. The second and third steps then performed a 3D search for tracks, respectively with and without the assumption that tracks originate from the beamline. Since misalignment of sensors, including a not complete closure of the vertex detector, leads to imperfect projections into *rz* space, the assumption of tracks pointing to the beamline in the first search for tracks was removed in

the new implementation to increase the track finding efficiency.

The new algorithm starts by searching for hits in at least four consecutive Velo R-sensors. A so-called seed track is created from these hits and is required to point approximately to the interaction region. Additional R-hits are searched for and, if sufficient hits are found, a search for Φ-hits is performed. A track is accepted if it consists of sufficient hits and has a good χ^2/DoF.

In the next step, hits that were not assigned to tracks by the previous step are used for a search for seeds which have hits in only three consecutive Velo R-sensors and point to the interaction region. The algorithm then proceeds, as in the previous step, by searching for additional R hits and Φ-hits before applying requirements on the quality of the track and marking hits as used.

The final step searches for seeds in unused Φ-hits without requiring that a seed points to the interaction region. Additional Φ- and R-hits are again searched for and good quality tracks are accepted.

3.2.2 Matching Velo Tracks to Muon Hits

Since the LHCb magnet provides a vertical dipole field, tracks are to a good approximation only deflected in the horizontal plane. Combining this with the high resolution of the measurement of a track's vertical direction by the Velo and the low hit occupancy of the muon stations gives an opportunity to identify muon tracks early in HLT1. This is done using a matching algorithm described below.

Since the vast majority of the particles created in the primary interaction are not muons, an early selection of muon candidates from available Velo tracks saves sufficient time to allow for looser p and p_{T} requirements in the subsequent forward pattern recognition step, which in turn improves the efficiency of the HLT1 muon selections.

Matching Algorithm

The matching algorithm searches for hits in the muon stations which can be associated to a Velo track. Hits in muon station M3 are considered first, as they are also required to be present in the offline muon identification. The preference for the M3 stations follows from the lower occupancy of the M3 station with respect to the M2 station, which speeds up the algorithm. The average occupancy of the muon stations in L0-accepted data is given in fig. 3.2.

As a first step, the algorithm defines a search-window in the M3 station. The vertical centre of the search-window is determined by extrapolating the Velo track to M3 in a straight line. Large amounts of material are present before and

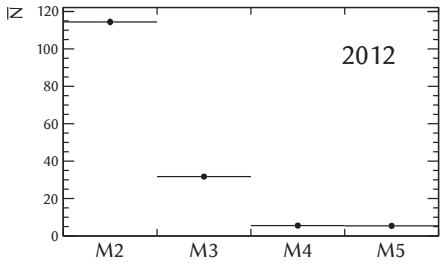


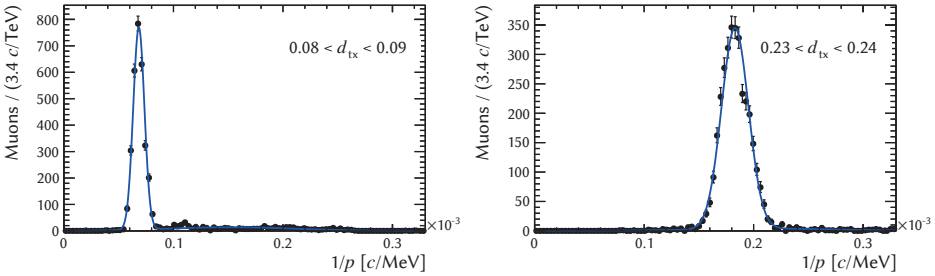
Figure 3.2: Average number of hits in muon stations M2 to M5 in L0-accepted events.

between the muon stations in the form of the calorimeters and layers of iron shielding, respectively. Multiple scattering in this material might cause a muon track to deviate from the track Velo extrapolation. To allow for multiple scattering, the vertical size of the search-window is chosen to be twice the height of a muon pad in the outermost region of M3. This ensures that a hit in a pad above or below the pad to which the Velo track is extrapolated will also be taken into account.

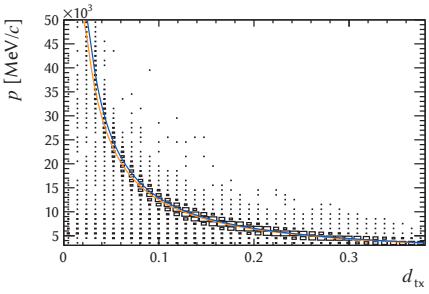
As the layers of iron shielding between the muon chambers absorb most muons with a momentum smaller than 6 GeV/c, the horizontal size of the search-window is obtained by calculating the maximum possible deflection of 6 GeV/c positive and negative particles. The deflection of a candidate is estimated by approximating the total effect of the magnet on the direction of a track as an instantaneous deflection at the focal plane of the magnet.

To obtain the relationship between the deflection, d_{tx}, and the offline reconstructed momentum, p, a fit of two Gaussians is performed to the distribution of the number of selected muons as a function of $1/p$ in bins of d_{tx} in a minimum bias data sample of identified muons; this is illustrated in fig. 3.3a for two bins in d_{tx}. A minimum bias data sample is suitable, since the misidentification rate for offline reconstructed muons is below 1%, as discussed in section 2.9. The first Gaussian describes the peak component, while the second takes the wide background component into account. The mean of the first Gaussian is extracted and shown in fig. 3.3c in a scatterplot versus the momentum, p. A straight line is fitted to the extracted means, which suggests that a parameterisation of the form:

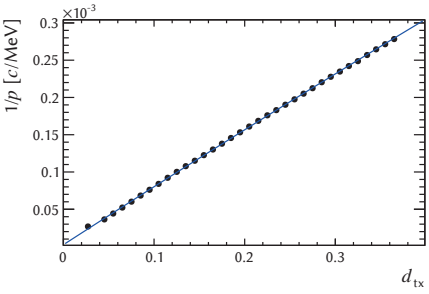
$$p = \frac{1}{a + b d_{\mathrm{tx}}}, \tag{3.3}$$



(a) Fit of two Gaussians to the distribution of the number of muons as a function of $1/p$ for two bins in d_{tx}. One Gaussian is clearly visible, while the other describes the small wide background component.



(b) 2D scatter-plot of the momentum of a muon versus its deflection by the magnet. The lines show the parameterisation used in 2011 and 2012 in blue and the one obtained from the fit shown in fig. 3.3c in orange.



(c) Central values of the peaks fitted to the distribution of the number of muons as a function of $1/p$ in bins of d_{tx}. A fit of a straight line is shown as the blue line.

Figure 3.3: Calibration of eq. (3.4) using offline reconstructed and identified muons.

is appropriate. The values of a and b extracted from the fit are:

$$\begin{aligned} \text{a:} \quad & (3.615 \ \ \pm 0.048 \ ) \times 10^{-6} \\ \text{b:} \quad & (7.6339 \pm 0.0035) \times 10^{-4} \end{aligned}$$

The resulting parameterisation is shown as the orange line in fig. 3.3b.

The matching algorithm was designed at the end of 2010 and at that time, the following parameterisation was obtained:

$$p = 175 + \frac{1255}{d_{\text{tx}}}, \tag{3.4}$$

which is shown as the blue line in fig. 3.3b, and which was used during 2011 and 2012 data taking. From the same figure it can be seen that while the orange and blue lines lie close together, they are different. The difference results in a 2 cm, or 0.5%, smaller search window for the parameterisation of eq. (3.3) compared to that of eq. (3.4). This means that although the new parameterisation better describes the data, the efficiency of the matching is only marginally affected.

The search for hits in the muon stations is illustrated in fig. 3.4, and proceeds by defining a search window in station M3. From the literature [55, sec. 5.5.2], it follows that the z-position of the magnet's focal plane, z_f, is a function of the horizontal direction of the velo track, t_x, parameterised by:

$$z_f = a + bt_x^2, \tag{3.5}$$

where a and b were found to be 1.069 m and 5.203 m, respectively. No dependence of z_f on the vertical direction of the velo track, t_y, was found. The x-coordinate, x_f, of the point on the candidate track at z_f is obtained by extrapolating the Velo segment in a straight line to z_f.

The left and right edges of the search-window are calculated from the deflection — assuming 6 GeV/c momentum — according to:

$$x_\mp = x_f + (z_{\text{M3}} - z_f)\tan(\theta_\mp) \mp x_w, \quad \text{where} \quad \tan(\theta_\mp) = \frac{t_x \mp d_{\text{tx}}}{1 \pm t_x d_{\text{tx}}}, \tag{3.6}$$

as illustrated in fig. 3.4. The minus sign refers to the edge at smallest x and the plus sign to the edge at largest x, z_{M3} is the z-coordinate of muon station M3 and x_w is an additional factor, which allows for multiple scattering in the horizontal direction. With the assumption of a minimal momentum of 6 GeV/c ($d_{\text{tx}} = 0.21$), and $x_w = 20$ cm, the search window is typically approximately 4.6 m wide, and covers roughly half of the width of muon station M3.

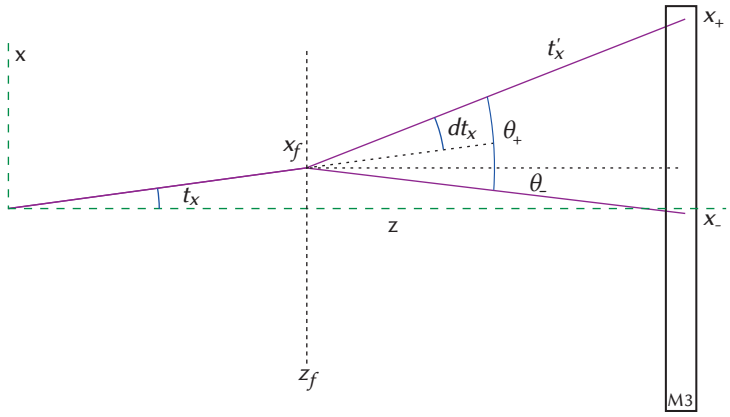


Figure 3.4: Illustration of the quantities involved in the calculation of the search window along the x-axis according to eq. (3.6)

A seed track is subsequently created for each hit inside the search-window by combining it with the Velo track segment. The direction of the track after the magnet is obtained from:

$$t'_x = \frac{x_h - x_f}{z_h - z_f}, \tag{3.7}$$

where the h subscript refers to the hit in M3 and the f subscript refers to the point at the focal plane of the magnet. Since d_{tx} is given by the difference between this slope, t'_x, and the slope of the Velo segment, t_x, the seed's momentum can be estimated using eq. (3.4).

The candidate track is then extrapolated to stations M2, M4 and M5 to search for additional hits inside a fixed-size search-window of 20×20 cm, which was optimised empirically. To maximise the efficiency of the search, a candidate track is provisionally accepted if it at least one additional hit is found.

In the final step of the algorithm, a linear least-squares straight-line fit in the horizontal plane is performed including the point at the focal plane of the magnet and any muon hits found. The uncertainty on the point at the focal plane of the magnet is obtained by extrapolating the uncertainty on the direction of the velo track segment. Half the size of the muon pads to which the muon hits correspond are taken as the uncertainties on the muon hits. The χ^2/DoF of the least-squares fit is required to be less than 25. As soon as a single candidate is found, the algorithm stops and the Velo track is accepted. Figure 3.5 shows the distributions of χ^2/DoF for candidates which correspond, or do not correspond, to an offline

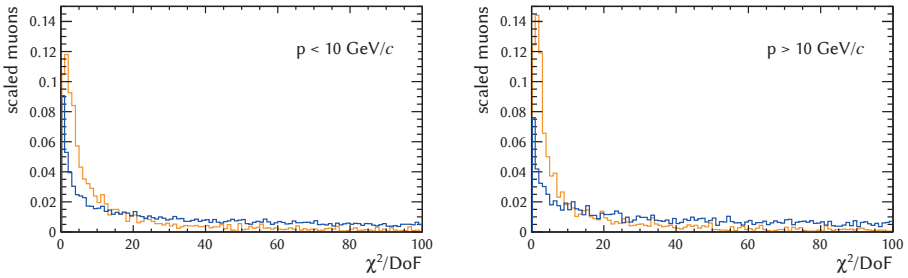


Figure 3.5: Normalised distributions of χ^2/DoF in 2012 minimum bias data for muon matching candidates with a Velo track segment that matches an offline reconstructed muon; the muon hits of the matching candidate are either matched (orange) or not matched (blue) to the offline reconstructed muon.

reconstructed and identified muon. Due to the long tail of the distribution, the requirement on χ^2/DoF is likely the main cause for the not-perfect efficiency of the matching algorithm. It can be seen that the efficiency and background rejection of the requirement are both better for higher momentum muons, which is also seen in fig. 3.7.

Performance of the Matching Algorithm

The rate at which the matching algorithm accepts muon candidates is important, since it determines the available CPU budget for subsequent operations. In 2011 and 2012, the rate was sufficiently low to allow track extension to the tracking stations, muon identification and track fit of all accepted segments, without the need of additional selection requirements on a high impact parameter with respect to any primary vertices. The probability of accepting Velo segments in L0 accepted events is shown in fig. 3.8; the average number of segments which are accepted per event is approximately 16 in 2012.

To evaluate the efficiency of the matching algorithm, a background subtracted sample of $B^+ \to J/\psi K^+$ decays is used, which was triggered independent of signal (TIS) with respect to any of the L0, HLT1 and HLT2 physics triggers. The TISTOS method for determining efficiencies has been described in more detail in section 3.1. The statistical background subtraction method using event weights will be described in section 4.2; in this case the event weights required for this procedure were calculated using a fit to the distribution of the dimuon J/ψ mass

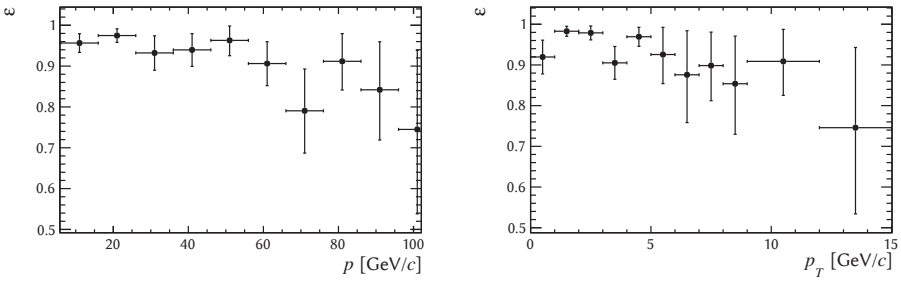


Figure 3.6: Left: efficiency of the HLT1 Velo pattern recognition algorithms for muons from $B^+ \to J/\psi K^+$ as a function of p in 2012. Right: efficiency as a function of p_T.

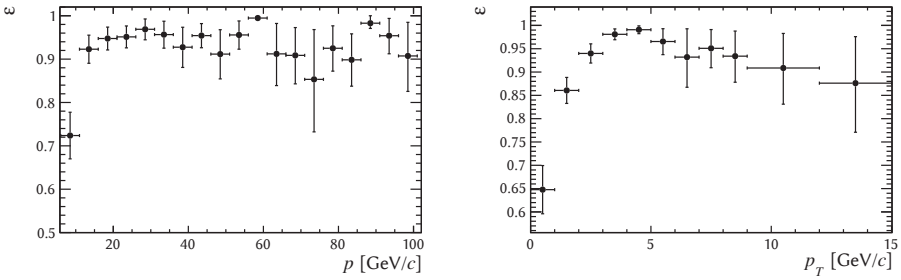


Figure 3.7: Left: efficiency of the matching algorithm for muons from $B^+ \to J/\psi K^+$, whose Velo track segments have been reconstructed by the Velo track finding, as a function of p in 2012. Right: efficiency as a function of p_T.

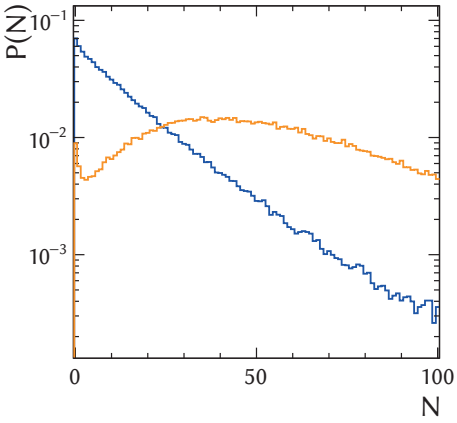
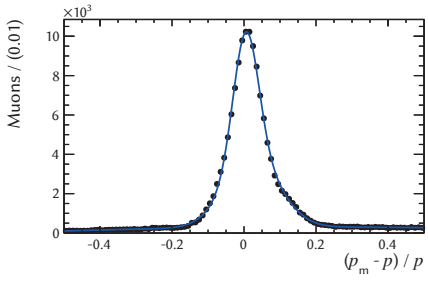


Figure 3.8: Probability of the matching algorithm accepting a number of Velo track segments, N, (blue) and the probability of N Velo segments being present in an event (orange) in a 2012 L0 accepted data sample.

peak. For each selected candidate, the Velo segments of the muon tracks are identified; only offline muons which have a Velo segment reconstructed in HLT1 are subsequently considered. The efficiency of the HLT1 Velo pattern recognition to find Velo segments of offline reconstructed and identified muons is shown in fig. 3.6, and has an average efficiency of $97 \pm 1\%$.

Each HLT1 Velo segment belonging to an offline reconstructed muon is used as input to the matching algorithm and the efficiency for accepting muons is determined; it is shown in fig. 3.7 as a function of muon p and p_{T}. For muons with a momentum above approximately 15 GeV/c, the efficiency is above 95%, while for low momentum muons the efficiency drops due to multiple scattering. The average efficiency is $93 \pm 1\%$.

The estimated momentum of the candidate can be used by the forward tracking step to speed up the search for hits in the tracking stations, as prior knowledge of the momentum can be used to define reduced search windows. To evaluate the quality of the momentum estimate, the relative difference between the estimated and offline reconstructed momentum is shown in fig. 3.9; A fit of the



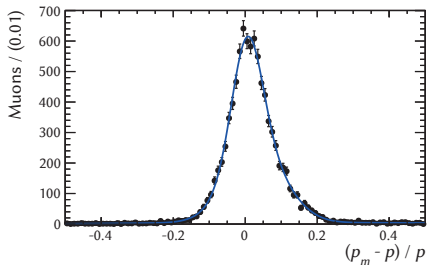
par.	value	
f_{G1}	0.478	$\pm$ 0.011
f_{peak}	0.8310	$\pm$ 0.0019
μ_1	0.005 16	$\pm$ 0.000 28
σ_1	0.022 37	$\pm$ 0.000 63
μ_2	0.035 08	$\pm$ 0.000 44
σ_2	0.075 52	$\pm$ 0.000 74
c_1	0.610	$\pm$ 0.022
c_2	-1.339	$\pm$ 0.093

Figure 3.9: Result of a fit to the distribution of the relative difference between the momentum estimated by the matching algorithm and the offline reconstructed momentum for offline reconstructed and identified muons with $p > 6$ GeV/c in 2012 L0 accepted data. Candidates are shown that have the same velo segment as the offline reconstructed candidate, and the smallest χ^2/DoF.

following PDF to the distribution gives the result listed in the same figure:

$$p' = (p_m - p)/p, \tag{3.8a}$$
$$\text{PDF}(p') = f_{\text{peak}}\left[(f_{\text{G1}}\text{G1}(p') + (1 - f_{\text{G1}})\text{G2}(p')\right] \tag{3.8b}$$
$$+ (1 - f_{\text{peak}})N_p(1 + c_1 p' + c_2 p'^2), \tag{3.8c}$$

where p_m is the momentum estimated by the matching algorithm, p is the offline reconstructed momentum, G1 and G2 are Gaussians and N_p is a normalisation constant. Equation (3.8a) gives the relative shift between the momentum estimated by the matching algorithm and the offline reconstructed momentum. Equation (3.8b) models the peak of muons with a correct momentum estimate, while eq. (3.8c) describes the combinatorial background contribution. The average width of the peaks is approximately 6%. The fit also shows, however, that if the region between -0.13 and 0.17 is considered good, 86% of the cases fall inside that region. The relatively high probability of wrongly estimating the momentum therefore results in a comparable loss of efficiency, which is the reason that the estimated momentum was not used as input to the forward pattern recognition step in 2011 and 2012.



par.	value	
f_{G1}	0.433	$\pm$ 0.043
f_{peak}	0.9252	$\pm$ 0.0032
μ_1	0.004 30	$\pm$ 0.001 28
σ_1	0.041 40	$\pm$ 0.001 83
μ_2	0.029 30	$\pm$ 0.002 04
σ_2	0.078 10	$\pm$ 0.001 94
c_1	1.000	$\pm$ 0.034
c_2	-1.973	$\pm$ 0.049

Figure 3.10: Result of a fit of a single Gaussian to the distribution of the relative difference between the momentum estimated by the matching algorithm and the offline reconstructed momentum for offline reconstructed and identified muons in 2012 L0 accepted data. The candidate created by the matching algorithm is required to have the same velo segment and muon with as offline reconstructed candidate.

Possible Improvements

There are several options for improving the performance of the matching algorithm in the future. The size of the search windows was empirically tuned using 2010 data and could be further optimised.

Figure 3.5 shows that the χ^2/DoF obtained from the fit of a straight line to the associated hits in the horizontal plane provides some discrimination between muons and background, but due to the long tails of the distribution for muons, its power is limited. For muons with a momentum above 10 GeV/c, the χ^2/DoF provides better discrimination, but a long tail is still present.

To improve this in the future, the uncertainties used in its calculation need to be calibrated, both for the point at the focal plane of the magnet and the muon hits. The fit of a straight line is currently performed only in the horizontal plane, but it could be extended with a fit in the vertical plane. The direction of the velo track would then serve as a constraint on the direction of the muon. Combining the calculated χ^2/DoF would likely improve the background rejection power of the fit. However, due to the small number of hits, and the approximations of a straight track without multiple-scattering, it is unlikely that the long tails in the χ^2/DoF could be completely removed by calibrating the uncertainties.

Alternatively, the combination of requiring hits in two muon stations with a relatively small search window size during the search for hits beyond M3 could

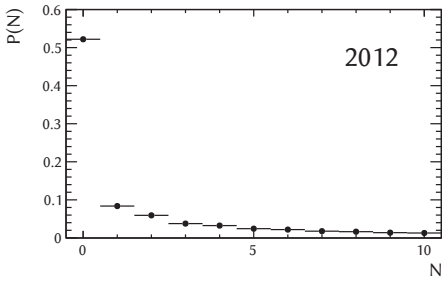


Figure 3.11: Probability of the matching algorithm creating a number of candidates, N, for a given velo segment in 2012 L0 accepted events.

be revisited. As discussed later, the muon identification applied at a later stage requires at least three hits. Imposing the same requirement in the matching algorithm would likely improve the purity of the sample at the cost of reduced efficiency. The efficiency could then be gained back by simultaneously increasing the size of the search window. In the current configuration there is only one more hit then there are degrees of freedom in the least-squares fit for a candidate that consists of the minimum number of two muon hits. Increasing the minimum number of required hits to three would therefore have the side-effect of improving the discriminating power of the least-squares fit.

The shape with tails in fig. 3.9 shows that while the estimated momentum is wrong too often for efficiency purposes, the precision is quite high when it is correct. Figure 3.10 instead shows the distribution of the relative difference between the momentum estimated by the matching algorithm and the offline reconstructed momentum for the "best" matching algorithm candidate: the candidate that has the same Velo segment as an offline reconstructed muon and the largest number of muon hits in common with it. Comparing fig. 3.10 to fig. 3.9, the background is observed to be smaller and the fraction of the peak increases from 83% to 93%, equal to the average efficiency of the matching algorithm. This implies that the the precision of the momentum estimate is indeed approximately 6%, and the smaller fraction of the peak in fig. 3.9 is a result of selecting wrong candidates based on the smallest χ^2/DoF in the horizontal plane.

In the current configuration of the matching algorithm, a relatively large number of candidates — 1.43 on average — is created due to the loose requirement of a single hit in addition to the seed hit in M3, as can be seen in fig. 3.11. To be able to benefit from possible a speed-up in the forward pattern recognition from the use of the momentum estimate, the number of matching candidates created

per Velo track segment must be reduced.

3.2.3 Primary Vertex Reconstruction

Once Velo tracks have been reconstructed, the reconstruction of primary vertices from these tracks proceeds in two steps: a search for vertex seeds and a fit of the seeds. The Velo tracks reconstructed in HLT1 are used as input for the seeding step, which calculates the distance of closest approach (DOCA) between all previously unused tracks and stores all combinations with DOCA < 300 μm and $\cos(\theta) < 0.999$, where θ is the opening angle between the tracks.

For each combination of two tracks the point corresponding to their closest approach to each other are stored and the position of the seed is calculated as the average of all points of closest approach. The two-track combinations are then looped over and only those which have a distance smaller than 3 mm to the position of the seed are kept. The mean is updated and the loop over combinations repeated until all included combinations are closer than 3 mm to their average position. If sufficient tracks remain part of the seed, its weighted mean is calculated for the z-position of the seed, where each track is assigned a weight of $\tan^2\theta$ where θ is the opening angle between the two tracks in the respective two-track combination, to take higher precision combinations into account. For the x- and y-position, the arithmetic mean is used.

To all seeds an adaptive fit is applied, which starts by collecting all tracks that have an impact parameter of at most 2 mm with respect to the seed. The remaining tracks are then iterated on, where each track is assigned a so-called Tukey weight of $\max(0, 1 - \chi^2_{\text{IP}}/c^2)$ where c is Tukey's constant, χ^2_{IP} is the difference in the χ^2 of the vertex between including and excluding the track. c starts at 15 for the first iteration and decreases to 3 for the final iteration. The fit iterates until the change in z-position of the vertex is less than 1 μm.

If the fit of a seed succeeds, the resulting vertex is added to the list of found vertices if it is sufficiently separated from previously found vertices and within a certain radius from the beamline. The tracks which belong to an accepted vertex are removed from the search for additional seeds. The algorithm repeats the seeding and fitting steps on unused tracks until no new vertices with more than five tracks are found.

3.2.4 Forward Track Finding

To determine the momentum of tracks, Velo tracks are used as a starting point for a search for hits in the inner and outer trackers by the so-called forward

track finding algorithm. The algorithm starts by opening a search window in the vertical (or x) layers of the T-stations by assuming a minimal momentum and transverse momentum and subsequently combining each hit inside the search window with the Velo track to create seed tracks.

Using the hit in the T-stations, the momentum of each track-seed is estimated. Track-parameterisations of all seeds are created using these estimated momenta, and used to calculated the x-coordinate of all seeds at a reference plane at fixed $z = 8520$ mm. x-coordinates obtained from hits that belong to the same track will cluster around the same x-coordinate on the reference plane and a search for hits becomes a search for clusters. This approach to pattern recognition is called a Hough Transformation [56]. Once a cluster of hits is identified, a candidate track is created, and its χ^2 is calculated. Hits with large contributions to the χ^2 are removed from the cluster.

If the calculated χ^2 for an xz-track candidate is sufficiently small, a search for stereo hits is performed. Using the direction of the Velo segment in the yz-plane, the corresponding x-coordinate of stereo hits is calculated and an analogous transformation to the reference plain and subsequent cluster search is performed to add appropriate stereo hits to the candidate track. Outliers that have large contributions to the χ^2 of a fit to the x-coordinates obtained from the stereo hits are removed.

Subsequently, a parabolic parameterisation is fit to the combined set of x-coordinates of hits in the vertical and stereo layers, and a straight line to the y-coordinates of the hits in stereo layers. Outliers are again removed according to their contribution to the χ^2 until the candidate track contains only hits with a contribution to the χ^2 less than 20. The difference in the y-coordinate at $z = z_{\text{ref}}$ between a straight line extrapolation of the Velo segment and the track candidate, Δy, is required to be smaller than:

$$\Delta y_{\text{max}} = 30 + \frac{300}{p^2}\text{mm}, \tag{3.9}$$

with p in GeV/c. Candidate tracks are rejected that consists of hits in fewer than nine different detector planes. After outlier removal, a quality variable, Q is defined as:

$$Q = \frac{5\Delta y}{\Delta y_{\text{max}}} + \frac{\chi^2/\text{DoF}}{10} + \frac{10}{|p|}, \tag{3.10}$$

with p in GeV/c. Small values of Q indicate good track candidates and the track candidate with the smallest value of Q, Q_{min}, is identified. Only candidates with $Q_{\text{min}} \leq Q < Q_{\text{min}} + 1$ are accepted. The number of hits, N_{hit}, on accepted track

candidates is counted and the candidate with the largest number of hits, N_{max}, is identified. Track candidates with $N_{\text{hit}} > 22$ or $N_{\text{max}} - 2 < N_{\text{hit}} < N_{\text{max}}$ are accepted.

Ghost tracks often result from the combination of a good Velo track segment with a good T-station track segment that, however, shares a significant number of hits with a T-station segment belonging to a different particle. To remove some of these combinations, the T-station hits of all tracks that have the same Velo segments are compared. If there is more than a 70% overlap between the T-station hits of two or more candidates tracks, only the best track in terms of Q is kept.

3.2.5 Muon Identification

The HLT1 dimuon selections use tracks that are identified as muons at two stages. The algorithm described in section 3.2.2 uses only Velo track segments and muon detector hits to identify muons candidates at the earliest possible stage. Its requirements are less stringent and it has a lower purity than the muon identification used at a later stage in HLT1, as well as in HLT2, which is identical to the offline binary identification criterion called IsMuon [57]. It tests the penetration of a track through the layers of iron shielding between the muon chambers. Once a track has been reconstructed, its trajectory is extrapolated to the z-coordinates of the muon stations and in each station hits are searched for in a field of interest (FOI). Depending on the momentum of the track, a different number of hits is required as tabulated in table 3.2.

momentum range	required stations
$3\,\text{GeV}/c < p < 6\,\text{GeV}/c$	M2 and M3
$6\,\text{GeV}/c < p < 10\,\text{GeV}/c$	M2 and M3 and (M4 or M5)
$10\,\text{GeV}/c < p$	M2 and M3 and M4 and M5

Table 3.2: Momentum ranges with hits in stations required for IsMuon.

3.2.6 Track Fitting

The LHCb track fitting algorithms fit a track model to a collection of hits using a Kalman filter based approach, which takes multiple scattering and energy loss into account [37]. The trajectory of a particle through the detector is parameterised by the track model as a collection of states at specific z-coordinates in the detector. A state is defined as $(x, y, t_x, t_y, q/pc)$, where $t_x = dx/dz$, $t_y = dy/dz$,

requirement	2011	2012
p	$> 6000\,\text{MeV}/c$	$> 3000\,\text{MeV}/c$
p_T	$> 500\,\text{MeV}/c$	$> 500\,\text{MeV}/c$

Table 3.3: Momentum requirements applied by the forward pattern recognition used in the HLT1 dimuon selections.

q is the charge of the particle and p its momentum. The forward design of the LHCb detector results in always increasing or always decreasing z-coordinates along the path of a particle through the detector. The fitting algorithm updates the states of the track model by adding information from each of the hits found by the pattern recognition algorithms one by one. At the z-coordinate of each hit, a state of the track is first predicted using information from previous states and then updated using the hit. The iteration over all the hits is first started from the hit with the smallest z-coordinate and then repeated starting from the hit with the largest z-coordinate. The result of these two iterations is averaged at each state to obtain the final set of track parameters and their uncertainties. As part of the fit, a number of hits with the largest contributions to the χ^2/DoF are removed.

In the HLT a simplified geometry is used to speed-up the determination of the amount of material a track traverses, which is needed to take multiple scattering into account.

3.2.7 HLT1 Dimuon Selections

HLT1 contains two dimuon selections, which start by searching for muon tracks by matching Velo track segments to at least two muon hits using the matching algorithm described in section 3.2.2. The rate at which Velo track segments are accepted is sufficiently low to provide the CPU budget for their extension to the tracking stations by the forward pattern recognition algorithm described in section 3.2.4, without the need for additional selection requirements on their impact parameter. The forward pattern recognition algorithm is configured with minimum (transverse) momentum thresholds listed in table 3.3. Successfully extended tracks are then extrapolated to the muon chambers and muon hits are searched for to identify them as muons according to the offline algorithm described in section 3.2.5. A simplified Kalman track fit is performed on tracks identified as muons followed by a selection requirement on the track's χ^2/DoF.

The first dimuon selection — called the high mass selection — creates dimuon

	requirement	high mass	low mass
tracks	$\chi^2_{\text{track}}/DoF$	< 3	< 3
	ISMUON	true	true
$J/\psi \to \mu^+\mu^-$	$\text{DOCA}_{\mu^+\mu^-}$	$< 200\,\mu m$	$< 200\,\mu m$
	χ^2_{vtx}/DoF	< 25	< 25
	$M(\mu^+\mu^-)$	$> 2.7\ \text{GeV}/c^2$	$> 1\ \text{GeV}/c^2$
	DLS	-	> 3

Table 3.4: Selection criteria applied by the HLT1 dimuon selections.

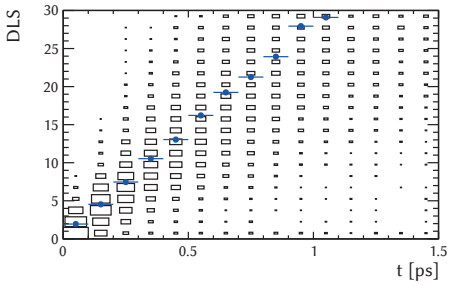
candidates from tracks identified as muons and requires them to form a vertex with a $\chi^2/\text{DoF} < 25$ and an invariant mass above 2.7 GeV/c^2. The second selection — called the low mass selection — selects dimuon combinations with an invariant mass above 1.0 GeV/c^2. The low-mass candidates are assigned the primary vertex to which they have the smallest impact parameter, and with respect to this primary vertex, they are required to have a decay-length significance (DLS) of $\text{DLS} > 3$. The DLS is defined by:

$$\begin{aligned} \mathbf{d} &= \mathbf{x}_{\mu\mu} - \mathbf{x}_{\text{PV}} \\ \text{DLS} &= \frac{|\mathbf{d}|}{\sigma_d}, \quad \text{where} \quad \sigma_d = \frac{1}{|\mathbf{d}|}\sqrt{\mathbf{d}^{\text{T}}(\mathbf{C}_{\mu\mu} + \mathbf{C}_{\text{PV}})\mathbf{d}}. \end{aligned} \tag{3.11}$$

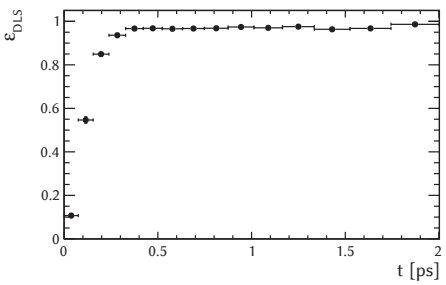
$\mathbf{x}_{\mu\mu}$ and $\mathbf{x}_{\text{PV}}$ are the coordinates of the $\mu^+\mu^-$ vertex and primary vertex, and $\mathbf{C}_{\mu\mu}$ and $\mathbf{C}_{\text{PV}}$ their respective covariance matrices. Table 3.4 summarises the selection requirements applied by the HLT1 dimuon selections.

For J/ψ particles originating from B-meson decays, the DLS of the J/ψ is highly correlated with the decay-time of the B-meson, as shown in fig. 3.12a. Below a DLS of approximately 10, the correlation is approximately linear, which makes it an appropriate proxy for the B-candidate decay-time in case only the muons have been reconstructed. From the figure it can be seen that the requirement of $\text{DLS}(J/\psi) > 3$ corresponds to a cut on the decay-time of approximately 0.1 ps. From fig. 3.12b, it can be seen that the requirement of $\text{DLS}(J/\psi) > 3$ results in a turn-on of the decay-time dependent efficiency up to approximately 0.4 ps.

Figure 3.13 shows the efficiency of the HLT1 dimuon high mass and low mass selections with respect to an offline selected sample of $B^+ \to J/\psi K^+$ candidates that are TIS at L0, HLT1 and HLT2. Since the mass of the dimuon candidates is above the threshold for the high mass selection, it performs better than the

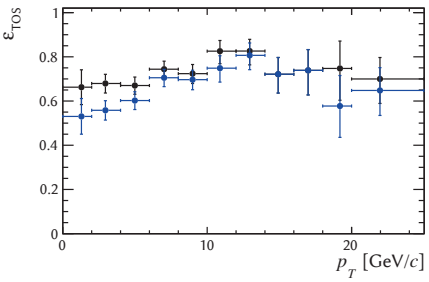


(a) J/ψ DLS versus B_s decay-time; the mean DLS is shown superimposed as the blue points.

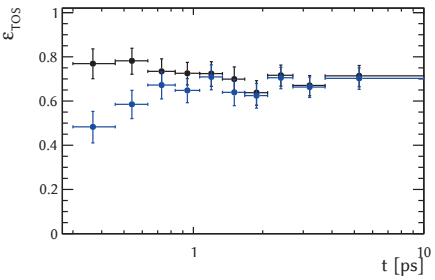


(b) Efficiency of the DLS(J/ψ) > 3 requirement as a function of the decay-time of the B_s.

Figure 3.12: J/ψ DLS of $B_s \to J/\psi\varphi$ signal decays simulated with 2012 conditions.

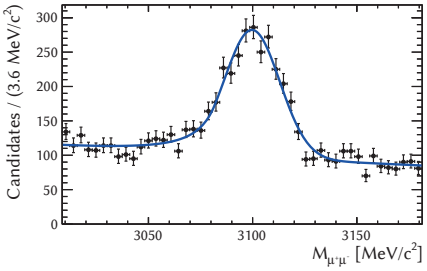


(a) Efficiency as a function of the p_T of the B^+.



(b) Efficiency as a function of the decay-time of the B^+.

Figure 3.13: Efficiency of the HLT1 dimuon selections for the combined 2011 and 2012 data sample of TIS triggered $B^\pm \to J/\psi(\to \mu^+\mu^-)K^\pm$ decays. The efficiency of the high mass dimuon selection is shown in black, while the low mass selection is shown in blue.



par.	value	
$N_{J/\psi}$	1869.0	$\pm$ 10.0
N_{bkg}	4700.0	$\pm$ 11.3
μ	3099.80	$\pm$ 0.56
σ	13.23	$\pm$ 0.57
α	1.37	$\pm$ 0.13
c	-0.0016	$\pm$ 0.0003

Figure 3.14: Result of a fit to the distribution of $M_{\mu^+\mu^-}$ of a sample of online reconstructed J/ψ candidates collected by LHCb during run 133584 on December 3rd 2012; the parameters are described in the text.

low mass selection. The average efficiency of the low mass selection is $65 \pm 2\%$, and in fig. 3.13b, the effect of the detachment requirement on the efficiency as a function of decay-time is clearly visible, as expected from fig. 3.12b (note the different scale on the horizontal axis). The average efficiency of the high mass selection is $72 \pm 2\%$ and although it is compatible with a uniform efficiency as a function of decay time, there is a hint of a decreasing trend for higher decay-time visible in fig. 3.13. This could be caused by a difference between the online and offline Velo pattern recognition leading to a reduced reconstruction efficiency at higher decay-times. This is discussed in more detail in section 5.1.

The combined output rate of the HLT1 dimuon selections is approximately 4 kHz or 5% of the total HLT1 output rate of approximately 80 kHz. A sample of online reconstructed and selected J/ψ candidates is shown in Figure 3.14, and the mass peak is clearly visible with a width that is similar to the width of the peak in the offline sample presented in chapter 6 (fig. 6.1). A Crystal Ball PDF [58, 59] was used to model the peak and an exponential PDF for the background. The Crystal Ball PDF is a Gaussian PDF with mean μ and width σ with a power-law tail ($n = 2$, in this case) on one side (the left side, in this case). The boundary between the power-law and Gaussian PDFs is defined by $(m - \mu)/\sigma > -\alpha$.

3.2.8 HLT1 Single Track Selections

Most heavy flavour selections require at least one final state track that is significantly detached from all primary vertices, and has a high momentum and transverse momentum. The single track HLT1 selections were designed to search for such tracks. There are three versions: an inclusive version that runs on all L0 accepted events, a muon version that only runs on L0 muon or dimuon accepted

events, and an electromagnetic version that only runs on L0 photon or electron accepted events.

The inclusive single-track selection runs on all events and is the primary trigger selection for beauty and charm decays to fully hadronic final states. To sufficiently reduce the number of Velo segments that are extended to the tracking stations by the forward pattern recognition (section 3.2.4) to fit in the available CPU budget, the inclusive single track selection starts with requirements on the number of hits and a minimal impact parameter with respect to any primary vertex of Velo segments.

At this point, the HLT configurations used in 2011 and 2012 differ. HLT configurations used in 2011 do not use the TT detector, and all accepted Velo segments are extended to the tracking stations. In the 2012 configurations, accepted Velo segments are extrapolated to the TT detector, and at least three associated TT hits are required for tracks that fall within the TT acceptance, while tracks that fall outside the TT acceptance are always accepted.

Thus selected Velo segments are subsequently extended to the tracking stations, followed by requirements on the number of hits in the tracking stations and a minimal momentum and transverse momentum. The simplified track fit described in section 3.2.6 is then performed, followed by a stringent requirement on the track χ^2/DoF to reject ghost tracks and a requirement on the track χ^2_{IP}.

The single-track muon selection proceeds in an analogous way, but only runs after a positive L0 muon or dimuon decision. It requires the Velo segment to be matched to at least 2 muon hits by the Velo-segment-to-muon-matching-algorithm described in section 3.2.2. Finally, several requirements are loosened with respect to the inclusive single track selection, such as the minimum momentum and transverse momentum required by the forward track finding algorithm and the maximum χ^2/DoF required for tracks. This is possible since tracks are required to be identified as muons by the offline identification algorithm.

The electromagnetic single track selection starts after a positive L0 photon or electron decision and includes loosened requirements with respect to the inclusive single track selection. Tables 3.5 and 3.6 summarise the requirements of all the single track selections in 2011 and 2012. The looser requirements in 2012 are a consequence of the increased CPU budget available due to the implementation of the deferred HLT and the inclusion of the TT validation for the inclusive and electromagnetic single track selections.

Figure 3.15 shows the efficiency of the HLT1 inclusive and single track muon selections for a sample of offline selected $B^+ \to J/\psi K^+$ candidates. The single

requirement	inclusive	muons	EM
N_{μ}	-	> 1	-
N_{Velo}	> 9	-	> 9
$N_{\text{Velo,miss}}$	< 3	-	< 4
IP	$> 100\ \mu m$	$> 100\ \mu m$	$> 100\ \mu m$
N_{TT}	-	-	-
$N_{\text{IT,OT}}$	> 16	-	> 15
p_{T}	> 1.7 GeV/c	> 1 GeV/c	> 1.2 GeV/c
p	> 10 GeV/c	> 8 GeV/c	> 6 GeV/c
$\chi^2_{\text{track}}/DoF$	< 2	< 2.5	< 2
χ^2_{IP}	> 16	> 16	> 16
ISMUON	-	true	-

Table 3.5: Selection criteria applied by the HLT1 single track selections in 2011.

requirement	inclusive	muons	EM
N_{TT}	> 3	-	> 3
p	> 1.6 GeV/c	> 1 GeV/c	> 1.2 GeV/c
p_{T}	> 3 GeV/c	> 3 GeV/c	> 3 GeV/c
$\chi^2_{\text{track}}/DoF$	< 2	< 2	< 2

Table 3.6: Selection criteria applied by the HLT1 single track selections in 2012 that are different than those in 2011.

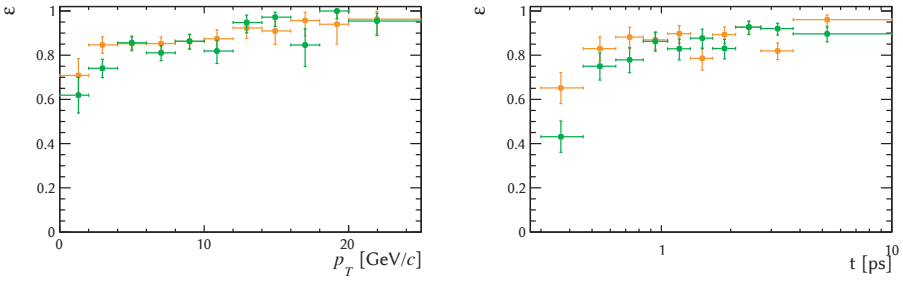


Figure 3.15: Efficiency of the HLT1 single track selections as a function of p_{T} and decay-time of the B^+ in the combined 2011 and 2012 data sample of TIS triggered $\mathrm{B}^{\pm} \to \mathrm{J}/\psi(\to \mu^+\mu^-)\mathrm{K}^{\pm}$ decays. The efficiency of the inclusive selection is shown in green, while the muon selection is shown in orange.

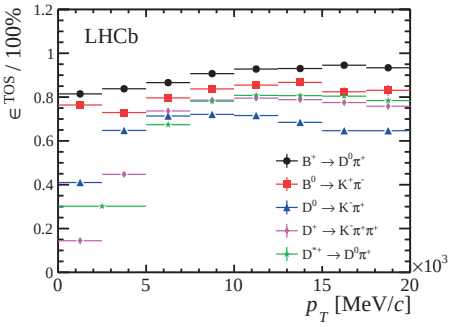


Figure 3.16: Efficiency of the HLT1 single track selections as a function of p_{T} of B or D mother particles for a number of offline selected heavy flavour decays; reproduced from [60].

track selections have the highest efficiency of all HLT1 selections, but do have a non-trivial decay-time dependence, due to the requirements on the χ^2_{IP} and the minimal number of Velo hits. Figure 3.16 shows the efficiency for the inclusive single track selection for heavy flavour decays to fully hadronic final states, for which this selection is the only available HLT1 selection. For D decays, the inclusive single track selection has an average efficiency of approximately 60% [60], with a lower efficiency for D mesons with a low p_{T}.

The rate of the inclusive single track selection is approximately 40 kHz while the single track muon selection accepts events at a rate of approximately 8 kHz, with some overlap with the inclusive single track selection and the dimuon selections.

3.3 HLT2

The second HLT stage, HLT2, consists of a reconstruction of all tracks that traverse the magnet, followed by inclusive and exclusive selections with additional on-demand reconstruction and particle identification. All reconstruction and particle identification algorithms are configured to perform as close as possible to their offline versions, whilst fitting in the HLT time budget.

The HLT2 inclusive selections are optimised to select heavy flavour decays based on general properties, without requiring all final state tracks to have been reconstructed; there are separate selections targeted at charm and beauty decays. These selections are the most efficient ones for final states without muons, and they are described in more detail in section 3.3.3.

Decays with muons in the final state, such as $\text{B}_\text{s} \to \mu^+\mu^-$ and $\text{B}_\text{s} \to \text{J}/\psi\,\varphi$, are most efficiently selected by one of the inclusive dimuon triggers. Due to the relatively low rate and high purity of dimuon events accepted by HLT1, the HLT2 inclusive dimuon triggers are rather simple; they are discussed in section 3.3.2. First, an overview of the reconstruction performed in HLT2 is given in section 3.3.1.

HLT2 additionally contains a number of so-called exclusive selections that target a specific decay of a beauty or charm hadron, and require reconstruction of all final state particles. Examples are: selections for decays of D^+ mesons to two and three body final states, a selection for decays of $\text{B}_\text{s} \to \text{K}^+\text{K}^-$ that avoids selection requirements on quantities that are correlated to the decay-time of the B_s, and selections targeted at B decays with photons in the final state, such as $\text{B}^0 \to \text{K}^*\gamma$.

3.3.1 Reconstruction in HLT2

At the start of HLT2, an additional search for Velo track segments is performed by the same algorithm which runs at the start of HLT1. In HLT1, only track-seeds that point to the interaction region and consist of hits in four consecutive Velo R-sensors are used. The additional search for Velo tracks in unused hits in HLT2 proceeds with a search for track-seeds that point to the interaction region and consist of hits in only three consecutive R-sensors. After the search for seeds consisting of three R-hits, a final search without a pointing constraint for seeds in Φ-hits is performed.

While additional Velo tracks are reconstructed at the start of HLT2, the search for primary vertices is neither repeated, nor are existing primary vertices updated if new tracks are found, to ensure that a consistent set of primary vertices is used in HLT1 and HLT2.

After the additional search for Velo tracks, the forward pattern recognition is now executed on all Velo track segments, independent of whether they were used for the forward reconstruction in HLT1. The requirements imposed on the minimal momentum and transverse momentum are lower than in HLT1; respectively 3 GeV/c and 500 MeV/c, in 2011, and 3 GeV/c and 300 MeV/c in 2012.

In 2011, the requirement of at least 500 MeV/c p_{T} significantly reduced the efficiency of the selections of charm hadron decays to more than two charged particles. To mitigate this efficiency loss, selections were constructed to first search for two-track combinations, followed by an additional forward pattern recognition step that searches for tracks with a minimal p_{T} of 300 MeV/c amongst unused hits. In 2012 sufficient CPU budget was available to set the p_{T} threshold in the forward pattern recognition to 300 MeV/c, which lead to a further improvement of the efficiency of the charm hadron selections.

HLT2 selections that include $\mathrm{K}^0_{\mathrm{S}}$ particles require special attention. The majority of the $\mathrm{K}^0_{\mathrm{S}}$ mesons decay outside of the Velo, which would lead to a low efficiency for such decays if the so-called downstream tracking is not executed. To create downstream tracks, a search for track segments in the T-stations is performed on hits that have not been assigned to any tracks by the forward pattern recognition. The T-station segments are used as a starting point for a search for hits in the TT detector. If sufficient TT hits are found and a candidate track is of good quality, it is accepted. The downstream tracks are then used as input to a $\mathrm{K}^0_{\mathrm{S}}$ selection algorithm and also for generic displaced vertex selections. The additional CPU budget available in 2012 allowed the time consuming search for track segments in the T-stations to be executed for all events.

Several HLT2 selections involve muons and the muon identification algorithm

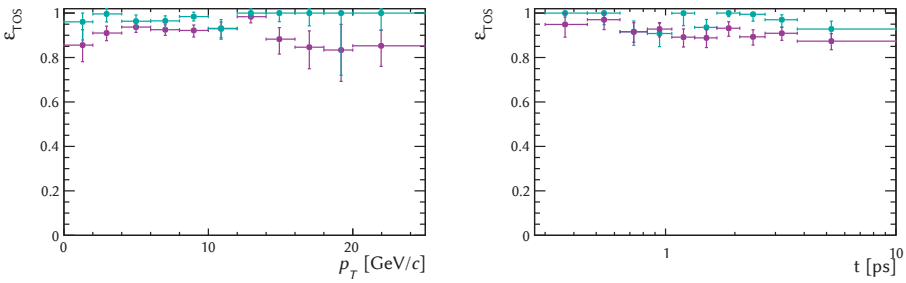


Figure 3.17: Efficiency of the HLT2 dimuon selections as a function of p_T and decay-time of the B^+ for the combined 2011 and 2012 data sample of TIS triggered $B^+ \to J/\psi K^+$ decays. The efficiency of the prescaled selection is shown in cyan, while the detached selection is shown in magenta. The effect of the prescale is not included in the shown efficiency of the prescaled selection.

is run on-demand for these selections. The HLT2 selections which include photons or electrons follow the same procedure for the calorimeter reconstruction.

3.3.2 HLT2 J/ψ Selections

At the HLT2 level, two trigger selections have been implemented to select J/ψ particles from B-decays: a detached selection that includes a requirement of $\mathrm{DLS}(J/\psi) > 3$ and a prescaled selection that does not include the requirement, but instead only runs on 20% of the input events. Both selections also require identified muons whose tracks have $\chi^2/\mathrm{DoF} < 5$, a dimuon vertex with $\chi^2/\mathrm{DoF} < 25$ and an invariant mass of $|M_{\mu^+\mu^-} - M_{J/\psi}| < 120$ MeV/c^2, where $M_{J/\psi}$ is the J/ψ mass of 3097 MeV/c^2 [14]. Requirements on the p_T of the muons have not been included to avoid explicitly biasing the angular distributions of e.g. $B_s \to J/\psi\varphi$.

The application of the DLS requirement decreases the rate of events accepted by the detached J/ψ selection by a factor 10, while the loss of efficiency is relatively small, about 1%. The efficiency of the requirement as a function of the decay-time of the B_s in simulated decays of $B_s \to J/\psi\varphi$ was shown in fig. 3.12b. The shape of the observed efficiency is very similar to what is found in data for $B^+ \to J/\psi K^+$, shown in fig. 3.17 and also to the efficiency obtained for $B_s \to J/\psi\varphi$, as described in chapter 5 (fig. 5.2).

The combined rate of all the HLT2 single and dimuon triggers is approximately 1 kHz or 20% of the total HLT2 output rate of 5 kHz in 2012.

3.3.3 Inclusive HLT2 Selections

The purpose of the inclusive HLT2 trigger selections is to accept events containing any partially reconstructed b-decays with maximum efficiency while providing sufficient background rejection to reduce the output rate to an acceptable level. An important advantage of the inclusive approach is that it is less sensitive to the reduced track finding efficiency of HLT2 compared to offline. Since, for the vast majority of selections, no particle identification is available for hadrons in HLT2, the inclusive selections focus on kinematic and geometrical properties of the partially reconstructed decay-vertices consisting of 2, 3 or 4-track combinations. They create n-body combinations from preselected input tracks which is followed by a multivariate selection on the n-body combinations, implemented as a boosted decision tree (BDT) [61].

There are three issues with the use of a boosted decision tree in the HLT. First of all, a limited set of simulated samples is available for training, which does not necessarily cover all interesting B-decays. This results in the risk that the BDT is trained to select B-decays only based on the specific properties of the available simulated samples, instead of using general properties of B-decays. Furthermore, the granularity at which the BDT splits its input observables should not be smaller than the experimental resolution or possible variation in time, i.e. due to changes in alignment or beam position. And finally, the algorithm which implements the BDT should be sufficiently fast to operate in the online environment.

To meet these conditions, a special type of boosted decision tree was designed that uses discretised input variables: the so-called bonsai boosted decision tree (BBDT) [62]. The discretisation is chosen such that the smallest difference between observables' values at which the BDT is allowed to split are larger than observed resolutions or variations. The discretisation allows the response of the BDT to be converted into lookup table, which makes the algorithm very fast and its CPU usage negligible in the HLT.

The BBDT uses the following input variables for n-body combinations, with $n \in \{2, 3, 4\}$:

- the sum of the transverse momentum of the final state particles, $\sum |p_{\mathrm{T}}|$,
- the smallest p_{T} of the final state particles, ${p_{\mathrm{T}}}^{\mathrm{min}}$,
- the invariant mass, m, of the n-body combination,
- the DOCA of the $n-1$ particle combination and the additional particle,
- the χ^2_{IP}, of the n-body combination,
- the flight distance (FD) χ^2 of the n-body combination,

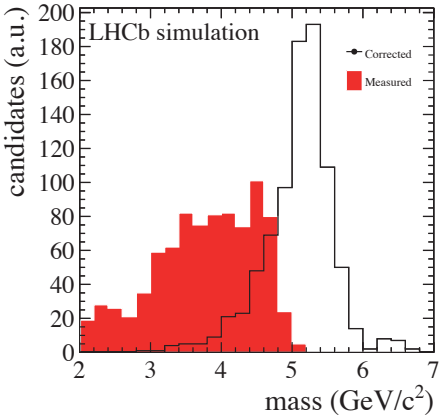


Figure 3.18: Reconstructed mass (red) and corrected mass (transparent) of $B^0 \to K^*\mu^+\mu^-$ where two out of the four final state tracks are included; reproduced from [60].

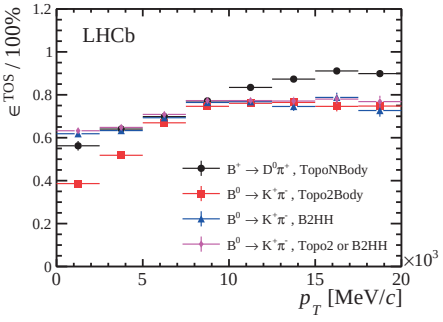


Figure 3.19: Efficiency of the HLT2 topological selections as a function of p_T; reproduced from [60].

- and the corrected mass, m_{corr}, which is defined as [63]:

$$m_{\text{corr}} = \sqrt{m^2 + |p_{\text{T,miss}}{}'|^2} + |p_{\text{T,miss}}{}'|, \tag{3.12}$$

  where $p_{\text{T,miss}}{}'$ is the missing momentum transverse to the direction of flight, which in turn is defined by the assigned PV and the n-body vertex.

Figure 3.18 shows both the invariant and corrected mass for $B^0 \to K^*\mu^+\mu^-$ decays when only two out of four final state tracks are included. The corrected mass has a much better resolution and peaks around the expected value whereas the 2-body invariant mass has a more uniform shape and is always below the expected value. Figure 3.19 shows the efficiency of the inclusive HLT2 selections with respect to offline selected $B^+ \to D^0\pi^+$ and $B^0 \to K^+\pi^-$ decays; the efficiency of the exclusive $B^0 \to K^+\pi^-$ HLT2 selection is also shown. The performance of the inclusive HLT2 selections for B decays is such that specific features or a proven increase in trigger efficiency are required to justify the additional CPU budget required for implementation of an exclusive selection. This is illustrated for $B^0 \to K^+\pi^-$ in fig. 3.19, where difference between the average efficiency of

the inclusive selection and the exclusive selection is only visible at low p_T of the B^0. The combined rate of the HLT2 inclusive selections is approximately 2 kHz or 40% of the total HLT2 output rate.

Events accepted first by the HLT1 dimuon and single track selections, and afterward by the HLT2 dimuon selections, are used as input for the measurement of ϕ_s from $B_s \to J/\psi\,\varphi$ decays, which is described in the following chapter.

Chapter 4

The Analysis of $B_s \to J/\psi\,\varphi$

The measurement of CP-violation in the interference between mixing and decay using the decay of $B_s \to J/\psi\,\varphi$ consists of a number of steps and ingredients. Before data can be analysed, the events of interest must have been triggered, reconstructed and selected. To be able to perform a fit of the observed decay-time and angular distributions, B-decays must be separated from background, and the theoretical PDF must be augmented to take experimental effects such as decay-time resolution and acceptance, and angular acceptance into account. The initial flavour of a candidate must also be determined, which is accomplished by a set of so-called flavour tagging algorithms. The maximum-likelihood fit to data is then performed to estimate the values and associated statistical uncertainties of the physics observables of interest, such as ϕ_s, Γ_s and $\Delta\Gamma_s$. All of these steps and ingredients are briefly introduced in this chapter; the decay-time acceptance and decay-time resolution are additionally discussed in greater detail in chapters 5 to 6. The parameter estimates, together with their systematic uncertainties, are given in chapter 7.

4.1 Reconstruction and Selection of $B_s \to J/\psi\,\varphi$ Candidates

The selection of $B_s \to J/\psi\,\varphi$ candidates commences with the combination of two tracks identified as muons and two tracks identified as kaons. A common vertex is fit to these four reconstructed tracks and from this combination, a $B_s \to J/\psi\,\varphi$ candidate is created. A primary vertex is subsequently assigned to the candidate according to the procedure described in the following section, and a kinematic

fit is performed to the combination of the candidate and its assigned primary vertex to extract properties, such as a candidate's decay-time and decay-angles. The steps in this procedure are described in more detail in the following sections.

4.1.1 Primary Vertex Assignment

The probability for the number of proton-proton interactions in a single bunch crossing is given by a Poisson distribution whose mean depends on the beam conditions. Some bunch crossings will result in more than one proton-proton interaction and therefore multiple reconstructable primary vertices. To calculate the decay-time of a $B_s \to J/\psi\,\varphi$ candidate its origin vertex must be known. In events which contain multiple primary vertices, a single vertex is selected by taking the vertex, with respect to which a $B_s \to J/\psi\,\varphi$ candidate has the smallest χ^2_{IP}. The impact parameter χ^2, or χ^2_{IP}, is the difference in χ^2 when the trajectory corresponding to the candidate is included respectively excluded from the primary vertex fit. It provides a measure of the how well a candidate "matches" a given primary vertex. If two or more primary vertices are located closely together, the wrong origin vertex can be assigned to a $B_s \to J/\psi\,\varphi$ candidate, resulting in a wrong decay-time, without affecting the decay angles.

4.1.2 Decay-Time

The decay-time is extracted by a kinematic fit of a model derived from the decay-tree to the reconstructed muons, kaons and assigned primary vertex by the decay-tree fitter (DTF) algorithm [64]. The DTF algorithm defines a χ^2 that includes constraint on the position of vertices, momentum conservation and decay-times as Lagrange multipliers, while measured tracks and the position of the primary vertex are included as measurement constraints. The decay-time constraint is introduced according to:

$$\tau = \frac{M}{|\mathbf{p}|^2}\,(\mathbf{x}\cdot\mathbf{p})\,, \tag{4.1}$$

where M is set to the world average of the B_s invariant mass [14], $\mathbf{p}$ is the 3-momentum of the particle and $\mathbf{x}$ is the flight-distance vector, i.e. the vector from the assigned primary vertex to the modelled decay-vertex of the candidate.

In the case of $B_s \to J/\psi\,\varphi$, the topology of the decay-tree results in the constraints that all of the tracks must originate from a common vertex and that the momentum and flight distance vectors should be parallel. An additional constraint on the invariant mass of the $\mu^+\mu^-$ can be applied, which improves the

resolution on the B_s mass. To avoid biasing the decay-time and decay-angles, this constraint is only applied when extracting the B_s mass.

The defined χ^2 is minimised using a Kalman filter approach, which has the important advantage that it deals more effectively with exact constraints and is computationally faster then a global χ^2 minimisation.

Equation (4.1) contains all sources of uncertainty which contribute to the decay-time uncertainty. If $\mathbf{x}$ and $\mathbf{p}$ are assumed to be parallel, it simplifies to the well known expression:

$$\tau = \frac{M\,|\mathbf{x}|}{|\mathbf{p}|}. \tag{4.2}$$

4.1.3 Mass PDF

The PDF used to describe the $J/\psi K^+K^-$ invariant mass distribution is a so-called Ipatia distribution [65], which was inspired by marginalisation of a Gaussian PDF with a width dependent on the estimated per-candidate mass uncertainty, σ_m, over the observed distribution of σ_m. This, under some assumptions on the distribution of σ_m, leads to the Ipatia distribution. To take final state radiation and the effect of the constraint of the J/ψ mass to the PDG value into account, exponential tails have been added to the left and right-hand side of the PDF, respectively. The PDF is given by:

$$I(m,\mu,\sigma,\lambda,\zeta,\beta,a_1,a_2,n_1,n_2) \propto \begin{cases} \frac{A}{(B+m-\mu)^{n_1}}, \text{if } m-\mu < -a_1\sigma \\ \frac{C}{(D+m-\mu)^{n_2}}, \text{if } m-\mu > a_2\sigma \\ \left((m-\mu)^2+\delta^2\right)^{\frac{1}{2}\lambda-\frac{1}{4}} e^{\beta(m-\mu)} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{(m-\mu)^2+\delta^2}\right), \text{otherwise,} \end{cases} \tag{4.3}$$

where $\delta \equiv \sigma\sqrt{\frac{\zeta\,K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}}$, $\alpha \equiv \frac{1}{\sigma}\sqrt{\frac{\zeta\,K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}}$, and A, B, C, D are obtained by imposing continuity and differentiability at $m-\mu < -a_1\sigma$ and $m-\mu = a_2\sigma$ and $K_\nu(z)$ is the modified Bessel function of the second kind. The parameter β is always fixed to zero.

The combinatorial background component is modelled with an exponential PDF.

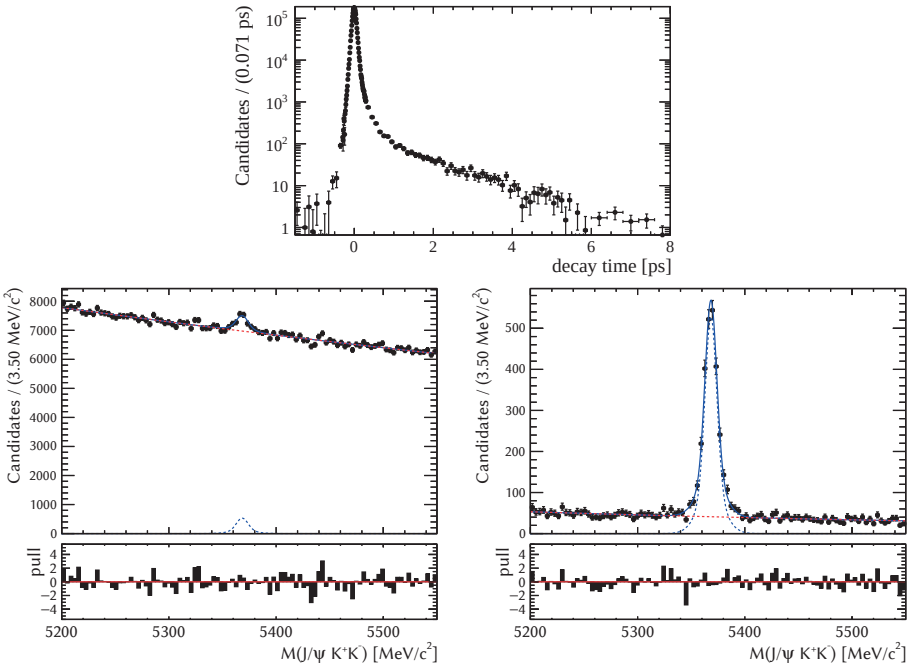


Figure 4.1: Candidates selected by the prescaled selection: decay-time distribution (top); $M_{\mathrm{J/\psi K^+K^-}}$ distribution, without (bottom left) and with (bottom right) application of a selection requirement on the decay-time of $t > 0.3$ ps.

parameter	value		
# background	4110.0	$\pm$ 7.3	
# B-decays	2479.0	$\pm$ 6.1	
$\overline{\sigma}$	7.51	$\pm$ 0.33	[MeV/c^2]
fraction 1st Gauss	0.67	$\pm$ 0.11	
mean mass	5368.38	$\pm$ 0.17	[MeV/c^2]
σ_σ	3.03	$\pm$ 0.62	[MeV/c^2]
background slope	$-0.001\,54$	$\pm$ 0.000 16	

(a) Sample selected with requirement of $t > 0.3$ ps.

parameter	value	
# background	696 056	$\pm$ 867
# B-decays	2491	$\pm$ 239
background slope	$6.52 \pm 0.12 \times 10^{-4}$	

(b) Sample selected without a requirement of $t > 0.3$ ps. All parameters not listed are fixed to the values listed in table 4.1a.

Table 4.1: Result of a fit to the distribution of $M_{J/\psi K^+K^-}$ for samples selected with and without a selection requirement of $t > 0.3$ ps.

4.1.4 Selection of $B_s \to J/\psi\,\varphi$ Candidates

By far most of the tracks originate from the primary vertexes, which results in a large number of combinatorial background candidates at low decay-times. This is demonstrated in fig. 4.1, which shows the decay-time distribution of a sample of selected candidates and the $M_{J/\psi K^+K^-}$ distribution with and without a requirement on the decay-time of $t > 0.3$ ps. Tables 4.1a and 4.1b list the corresponding fit results. The number of candidates at low decay times is too large to store. On the other hand, introducing additional tight selection criteria to reduce the number of events written to disk reduces the signal efficiency and could result in additional biases and acceptance effects. At the same time the number of background candidates decreases quickly with decay-time, while a B-decay candidate's contribution to the sensitivity of the analysis is roughly independent of decay-time for ϕ_s while for $\Delta\Gamma_s$ its contribution to the uncertainty decreases with the RMS. In the end a single decay-time cut of $t > 0.3$ ps has been chosen as a compromise between background rejection, signal retention and sensitivity.

The calibration of the decay-time resolution model, however, requires access to an "edge" of the decay-time distribution, for example at $t = 0$. To retain a sample that contains such events, its selection must not contain requirements

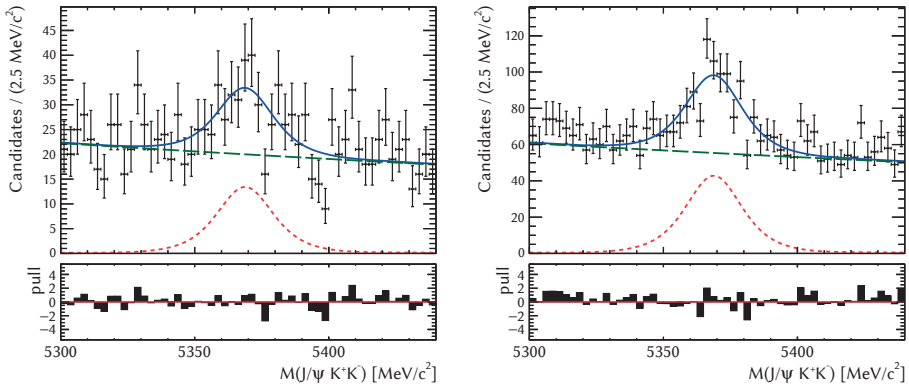


Figure 4.2: Fit to the $J/\psi K^+K^-$ invariant mass distribution of the 2011 (left) and 2012 (right) samples rejected by the single candidate selection.

on the decay-time. The final selection has, therefore, been split into two parts, similar to the HLT. The so-called prescaled selection does not include decay-time biasing requirements, but only approximately 20% of the events in which it selects a candidate are stored, while the so-called detached selection includes the requirement of $t > 0.3$ ps and all events in which it selects a candidate are stored. An additional advantage of this approach is that due to the high background rejection of the decay-time requirement, it has been possible to loosen other criteria to reduce implicit biases and acceptance effects. Table 4.2 lists the cuts applied as part of the prescaled and detached selections.

4.1.5 Multiple Candidates

To illustrate the possibility of encountering two signal decays in a single event we investigate two scenarios and assume a 100% selection efficiency: creating a $B_s \to J/\psi\varphi$ from both quarks in a produced $b\bar{b}$-pair and creating two $B_s \to J/\psi\varphi$ candidates in separate pp interactions in the same bunch-crossing. The branching fraction of $B_s \to J/\psi\varphi$ with $J/\psi \to \mu^+\mu^-$ and $\varphi \to K^+K^-$ is approximately 3×10^{-5}. Combining this with the fraction of b quarks that hadronise into B_s-mesons, $f_s \sim 0.1$ [14], the probability of one of the quarks from a $b\bar{b}$-pair decaying as $B_s \to J/\psi\varphi$ is $\sim 6 \times 10^{-6}$.

This probability is so small that the probability that both quarks produce a signal decay is ignored when the probability for creating multiple $B_s \to J/\psi\varphi$ decays in the same bunch-crossing is investigated. During typical LHCb 2012

decay mode	selection requirement	pre-selection	final selection
all tracks	$\chi^2_{track}/nDoF$	< 3	< 3
	clone distance	–	> 5000
$J/\psi \to \mu^+\mu^-$	-$\Delta\log\mathcal{L}(\mu\text{—}\pi)$	> 0	> 0
	$min(p_T(\mu^+), p_T(\mu^-))$	$> 0.5\,\mathrm{GeV}/c$	$> 0.5\,\mathrm{GeV}/c$
	$\chi^2_{vtx}/nDoF(J/\psi)$	< 16	< 16
	$\lvert M(\mu^+\mu^-) - M(J/\psi)\rvert$	$< 80\,\mathrm{MeV}/c^2$	$\in [3030,\ 3150]\,\mathrm{MeV}/c^2$
$\phi \to K^+K^-$	-$\Delta\log\mathcal{L}(K\text{—}\pi)$	> 0	> 0
	$p_T\,(\phi)$	$> 0.5\,\mathrm{GeV}/c$	$> 1\,\mathrm{GeV}/c$
	$M(\phi)$	$< 1050\,\mathrm{MeV}/c^2$	$\in [990, 1050]\,\mathrm{MeV}/c^2$
	$\chi^2_{vtx}/nDoF(\phi)$	< 25	< 16
$B_s \to J/\psi\,\varphi$	$M(B_s)$	$\in [5150,\ 5550]\,\mathrm{MeV}/c^2$	$\in [5200,\ 5550]\,\mathrm{MeV}/c^2$
	$\chi^2_{vtx}/nDoF(B_s)$	< 20	< 10
	$\chi^2_{DTF(B+PV)}/nDoF(B_s)$	–	< 5
	$\chi^2_{IP}(B_s)$	–	< 25
	$\chi_{IP,next}(B_s)$	–	> 50
	$t(*)$	> 0.2ps	$[0.3, 14.0]$ps

Table 4.2: Selection criteria for $B_s \to J/\psi\,\varphi$ candidates applied during the pre-selection and and final selection steps. The clone distance is the Kullback-Liebler Distance as defined in [66], -$\Delta\log\mathcal{L}$ is the log-likelihood difference between two particle hypotheses and $\chi^2_{DTF(B+PV)}$ is the χ^2 provided by the kinematic fit to the muon and kaon tracks and the position of the primary vertex. The χ^2_{IP} was described in section 4.1.1 and if multiple primary vertices have been reconstructed in an event, the $\chi_{IP,next}$ is defined as the second smallest χ^2_{IP} of a particle with respect to each of the primary vertices. (*) A slightly different algorithm is used to calculate the decay-time in the pre-selection and final selections. The difference of 0.1 ps between the decay-time requirements in the pre-selection and final selection has been shown to cover all differences.

running conditions, the instantaneous luminosity in the LHCb interaction point is 4×10^{32} cm^{-2}s^{-1} with 1262 pairs of colliding bunches [44]. This leads to a luminosity per bunch pair of $\mathcal{L}_{bb} = 3.17^{29}$. Given the revolution frequency, f_{rev}, of the LHC of 11245 Hz [67], a B_s production cross-section of $\sigma(pp \to B_s X) = 10.5 \pm 1.3$ µb [68] results in an average number of B_s mesons created per crossing of $\mu = \sigma_{B_s} \mathcal{L}_{bb} / f_{rev} = 2.8 \times 10^{-4}$. Given the Poisson distribution of the number of observed B_s-mesons per crossing, the probability of observing at least 2 B_s-mesons in a beam-beam crossing is then approximately 4×10^{-8}, without taking subsequent branching fractions into consideration.

The probability of observing two signal decays in a single event is therefore negligible, which implies that if multiple candidates are accepted by the final selection, at most one is a signal decay. Considering the case where a signal decay is present, there are two scenarios for accepting multiple candidates. If one or more muon or kaon tracks from another source than the signal decay of interest are close in phase space to one of the tracks from the signal decay, multiple candidates may pass the final selection, while only one corresponds to the true candidate. The multiple candidates are partial clones in this case. In the other scenario, the event contains a background candidate from the same or another *pp*-interaction, where the background and signal candidates do not share any tracks.

In case of multiple candidates, a single candidate is chosen, based on an observable that is uncorrelated to the decay-time and decay-angles. For this, the χ^2_{DTF} obtained from the kinematic fit described in section 4.1.2 is used: the candidate with the lowest lowest χ^2_{DTF} is chosen.

To estimate the number of multiple candidates in the sample before the single candidate selection, a simultaneous fit is performed to the $J/\psi K^+K^-$ invariant mass distribution of the combined 2011 and 2012 samples. The mass PDF is the same PDF as is used to obtain sWeights to extract the CP phase ϕ_s. The result of the fit is shown in table 4.3 and for those samples whose candidates do not pass the single candidate selection in fig. 4.2. The fact that a peak is still present in fig. 4.2 shows that some signal candidates are rejected by the multiple candidate selection requirement; the number of signal candidates in the rejected sample is approximately 0.4% of the total number of signal candidates. The signal-to-background ratio in the rejected samples is, however, much smaller than in the accepted samples, which provides additional motivation for the inclusion of the single candidate selection requirement.

parameter	value	
# bkg_{2011}	$38\,866.0 \pm 21.3$	
# bkg_{2012}	$97\,844.0 \pm 34.0$	
# $bkg_{2011,mult}$	2968.0 ± 5.7	
# $bkg_{2012,mult}$	8238.0 ± 9.5	
# sig_{2011}	$29\,950.0 \pm 19.1$	
# sig_{2012}	$66\,013.0 \pm 28.9$	
# $sig_{2011,mult}$	119.0 ± 2.0	
# $sig_{2012,mult}$	375.0 ± 3.4	
α_{mass}	$-1.500 \pm 0.026 \times 10^{-3}$	
μ_{mass}	5368.261 ± 0.027	$[\,MeV/c^2\,]$
σ_{mass}	8.347 ± 0.033	$[\,MeV/c^2\,]$

Table 4.3: Results of a simultaneous fit to the $J/\psi K^+K^-$ invariant mass spectrum, where the sample has been split between periods 2011 and 2012 and also between candidates which do or do not fulfil the single candidate selection requirement. Parameters labelled with "mult" represent the sample rejected by the single candidate selection requirement.

4.1.6 Trigger Selections

To maximise the selected sample size, several HLT selections are combined. Due to its high efficiency and high purity, the detached HLT2 dimuon selection, described in section 3.3.2, is sufficient at the HLT2 level. The selection requirement that $DLS(J/\psi) > 3$ results in a decay-time acceptance, since $DLS(J/\psi)$ is correlated with the decay-time of the $B_s \to J/\psi\,\varphi$ candidate.

The HLT1 high-mass dimuon selection described in section 3.2.7 is used because it has an efficiency of approximately 70%, utilising only selection requirements which do not result in a decay-time dependent acceptance. To increase the sample size, the HLT1 single track and HLT1 single track muon selections are included, which results in approximately 15% extra signal candidates. Since the HLT1 single track selections include several requirements on decay-time correlated observables, the decay-time acceptance for candidates which have only been selected by these two HLT1 selections is non-trivial.

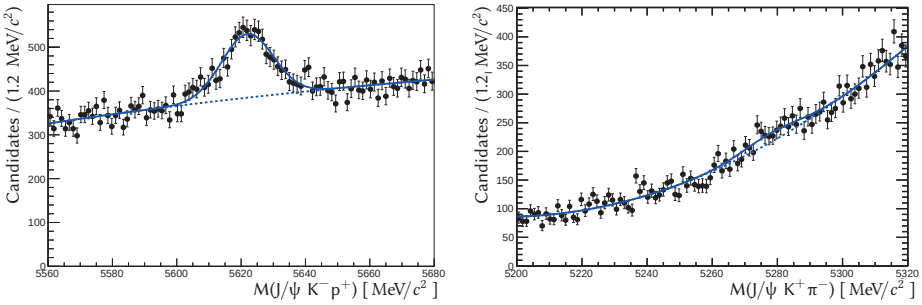


Figure 4.3: Distributions of $M_{\mathrm{J/\psi K^+K^-}}$ candidates in the side bands with the proton mass (bottom left) or pion mass (bottom right) assigned to either kaon (reproduced from [69]).

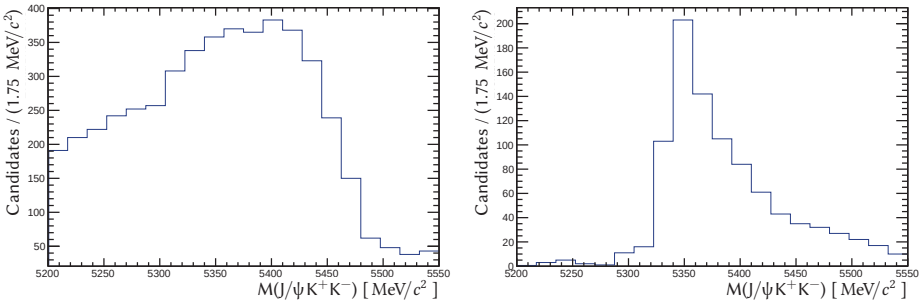


Figure 4.4: Distributions of $M_{\mathrm{J/\psi K^+K^-}}$ for the $\Lambda_b \to \mathrm{J/\psi p}K$ (left) and $\mathrm{B^0} \to \mathrm{J/\psi K^*}$ (right) simulated samples, reconstructed as $\mathrm{B_s} \to \mathrm{J/\psi}\varphi$ (reproduced from [69]).

4.1.7 Reflection Backgrounds

Studies of the data in the side-bands of the $M_{J/\psi K^+K^-}$ distribution have shown that in addition to B-decays and combinatorial background, the distribution also contains a reflection background component. Candidates from a decay of $\Lambda_b \to J/\psi p K$, with a misidentified proton, and candidates from a decay of $B^0 \to J/\psi K^*$, with a misidentified pion, that are not rejected by the final selection, have an $M_{J/\psi K^+K^-}$ distribution with a component under the B_s mass peak. In addition, the shape of these distributions under the B_s mass peak cannot be obtained by interpolating the shape in the sidebands.

The size of the reflection background components is measured in the side-bands by assigning either kaon the proton or the pion mass and performing a fit to the invariant mass distribution. To estimate the yield of $\Lambda_b \to J/\psi p K$, both side-bands have been used ($M_{J/\psi K^+K^-} < 5340$ MeV/c^2 or $M_{J/\psi K^+K^-} > 5400$), while for $B^0 \to J/\psi K^*$ only the higher side-band is used. The results of such fits is shown in fig. 4.3.

Once the yield of the reflection background components is known in the side-bands, it is extrapolated to the full mass range using the $M_{J/\psi K^+K^-}$ distributions for $\Lambda_b \to J/\psi p K$ and $B^0 \to J/\psi K^*$ reconstructed as $B_s \to J/\psi \varphi$, which have been obtained from simulation and are shown in fig. 4.4.

4.2 Statistical Background Subtraction

The sample selected by the final selection still contains some background, which can either be subtracted or included in the observable PDF. For the analysis described here, a statistical background subtraction technique known as sWeights [70] is applied. The technique uses the result of a fit to the distribution of one, or several, of the observables to assign a weight to each candidate, such that the weighted data sample effectively becomes that of one of its components. The weights can only be computed from those observables that are not correlated with any of the remaining observables of interest. For this analysis, the $J/\psi K^+K^-$ invariant mass is used to calculate the sWeights.

The presence of reflection backgrounds pose a problem for the use of sWeights, because the misidentification of one of the final state particles introduces a correlation between the invariant mass of the reflection candidate and its decay angles. The reflection backgrounds therefore need to be taken into account before the fit to $M_{J/\psi K^+K^-}$ is performed to determine the sWeights. This is achieved by injecting a number of simulated candidates corresponding to the identified

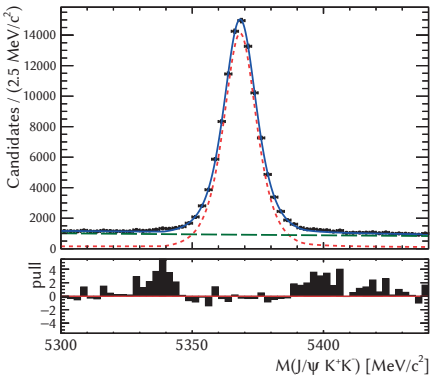


Figure 4.5: Fit to the $J/\psi K^+K^-$ invariant mass distribution of combined 2011 and 2012 samples which have passed the final selection; no reflection background subtraction was performed.

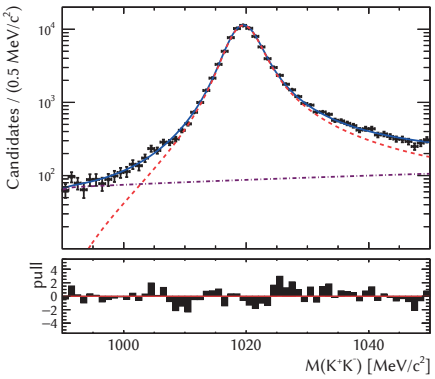


Figure 4.6: Result of a fit to the M_{KK} distribution of selected B decay candidates in the combined 2011 and 2012 samples.

sources and equal to the respective extrapolated yields into the sample, with negative weights. Since the decays that result in the reflections are explicitly known, candidates simulated from these decays are known to correctly reproduce the required distributions of decay-time and decay-angles. The tagging and angular distributions of the simulated samples used for injection of the reflection backgrounds are reweighed to approach those measured in data for the appropriate decays.

Figure 4.5 shows the $J/\psi K^+K^-$ invariant mass distribution of the combined 2011 and 2012 sample of selected candidates, to which the PDF described in section 4.1.3 has been fit. The result of a fit to the $J/\psi K^+K^-$ invariant mass is used to calculate the sWeights for the fit to the decay-time and angular distributions.

A potential difficulty in using the calculated events weights in a subsequent maximum-likelihood fit is that uncertainties will be estimated wrongly [71, p. 172, Ch. 8]. To estimate the size of this effect, a set of simplified simulations (toys) may be used, where observable distributions are generated according to the shapes fitted in data, and then fitted again to check if the resulting parameters agree with those used in the generation. Such toys provide a check whether the central value of an input parameter is on average recovered and whether the uncertainty of the recovered values matches the observed statistical fluctuations. To

evaluate this, the so-called pull is plotted, which is defined as $(p_{\text{obs}} - p_{\text{gen}})/\sigma_p$. The distribution of the pulls should be a Gaussian centred around 0 with a width of 1.

To estimate the effect of the use of sWeights on the uncertainties of fitted parameters, the fitting procedure for the simplified simulations has been adapted to consist of a first fit to the appropriate invariant mass distribution, followed by a calculation of the sWeights and a final fit to the weighted distributions. The results of such a test performed for the use of sWeights as part of while the decay-time resolution model calibration is described in appendix C.

4.3 S-Wave and K^+K^- Invariant Mass Distribution

As described in section 1.2.1, the $J/\psi K^+K^-$ final state contains a contribution which is non-resonant in the K^+K^- invariant mass M_{KK}. To show the presence of this non-resonant component, a fit to the M_{KK} distribution of selected $B_s \to J/\psi\,\varphi$ candidates is performed, where sWeights have been used to subtract the non-$B_s \to J/\psi K^+K^-$ contribution. The φ component is modelled by a relativistic Breit-Wigner PDF [72], which is given by:

$$\text{PDF}(M) = \frac{k}{\left(M^2 - M_\varphi^2\right)^2 + M_\varphi^2\Gamma^2}, \tag{4.4}$$

$$k = \frac{2\sqrt{2}M\Gamma\gamma}{\pi\sqrt{M^2+\gamma}}, \tag{4.5}$$

$$\gamma = \sqrt{M^2\left(M^2+\Gamma^2\right)}, \tag{4.6}$$

where M is the observed K^+K^- invariant mass, M_φ the pole-mass of the φ, and Γ its decay-width. The non-resonant component is described by an empirical model defined by:

$$\text{PDF}_{\text{em}}(M) \propto (1 - e^{-(M-M_0)/c_1})\frac{M}{M_0} + c_0\left(\frac{M}{M_0} - 1\right), \tag{4.7}$$

where $M_0 = 987.36$ is twice the kaon rest mass. The mean and width of the relativistic Breit-Wigner are fixed to their world average values, of 1019.46 MeV/c^2 and 4.27 MeV/c, respectively [14]. Both PDFs are convoluted with a Gaussian PDF with mean μ and width σ, $G(m)$, to model the K^+K^- invariant mass resolution. The full PDF is given by:

$$\text{PDF}(M) = G(M) \otimes (f_\varphi \text{BW}(M) + (1 - f_\varphi)\text{PDF}_{\text{em}}(M)), \tag{4.8}$$

parameter	value
bkg: c_0	8.2 ± 2.1
bkg: c_1	0.65 ± 0.59
f_ϕ	0.945 ± 0.003
μ	0.440 ± 0.012
σ	1.23 ± 0.02

Table 4.4: Result of a fit to the M_{KK} distribution of signal candidates.

where $\mathrm{BW}(M)$ is given by eq. (4.6). The result of a fit of this PDF to the background-subtracted distribution of M_{KK} for $\mathrm{B_s} \to \mathrm{J/\psi\,\varphi}$ signal candidates is listed in table 4.4 and shown in fig. 4.6; the fraction of the non-resonant purple dashed-dotted line is observed to be approximately 4% in this fit.

Since the line-shape of this non-resonant component is not well known, the decision was made to perform the final fit in bins of M_{KK}. The line-shapes of the non-resonant and resonant components depend on M_{KK}, which results in M_{KK}-dependent interference terms in the products of the three polarisation amplitudes of the P-wave, $\mathcal{A}_0$, $\mathcal{A}_\perp$ and $\mathcal{A}_\parallel$, with that of the S-wave, $\mathcal{A}_S$, which was ignored in eq. (1.24).

A correction factor is required when integrating over a certain M_{KK} range, as $\langle ps^* \rangle \neq \langle p \rangle \langle s^* \rangle$; where * denotes complex conjugation, p denotes the P-wave amplitudes, and s the S-wave amplitude. This correction is applied by introducing the factor C_{SP} in the products of the P- and S-wave amplitudes in the appropriate integrals. It is calculated by integrating the product of the normalised line-shapes for p and s^* over the M_{KK} range of the bin. The $f_0(980)$ is assumed to dominate the S-wave line-shape and is modelled by a Flatté function [73], and the P-wave line-shape is modelled by a relativistic Breit-Wigner function (eq. (4.6)). C_{SP} is then given by:

$$\int_{m^L}^{m^H} dM_{(\mathrm{K^+K^-})} p \times s^* = C_{SP} \mathrm{e}^{-i\theta_{SP}}. \tag{4.9}$$

To obtain the final C_{SP} factors, the M_{KK} resolution needs to be taken into account, which is achieved by treating the binning in the reconstructed M_{KK} as an additional selection efficiency on the true M_{KK} that appears in eq. (4.9) [69]. This efficiency is determined from a sample of simulated and selected $\mathrm{B_s} \to \mathrm{J/\psi\,\varphi}$ candidates and eq. (4.9) is modified appropriately to obtain the final C_{SP} factors listed in table 4.5.

M_{KK} range (MeV/c^2)	C_{SP}
990—1008	0.9178
1008—1016	0.9022
1016—1020	0.8619
1020—1024	0.8875
1024—1032	0.9360
1032—1050	0.9641

Table 4.5: C_{SP} factors obtained using a modified version of eq. (4.9) to take the M_{KK} resolution into account; reproduced from [69].

The sWeights used in the final fit are calculated for each bin in M_{KK} separately to take changes in the B decay and background yields across the bins into account. The PDF used for the calculation of the sWeights is a simultaneous PDF for the $\mathrm{J/\psi K^+K^-}$ invariant mass. The yields of the B decay and background components and the slope of the background component are separate parameters for each bin in M_{KK}, while the parameters which determine the shape of the B decay peak are shared between the bins.

Distributions of decay-time and helicity angles for B decay candidates after statistical and reflection background subtraction are shown in fig. 4.7.

4.4 Decay-Time Acceptance

To avoid decay-time acceptances, the selection requirements on the decay-time, which are part of the pre-selection and final selection, are aligned to the range in which the final fit is performed. The strictest requirement of $t > 0.3$ is applied as part the final selection, which is thus the lower boundary of the range in which the final fit is performed.

The first source of a decay-time dependent acceptance is the result of a decreased efficiency of the VELO track finding algorithms for tracks with larger impact parameter to the beamline. During the search for Velo tracks, the assumption is made that given two track candidates, which share hits, the candidate with the smaller impact parameter to the beamline is the better candidate.

Since tracks from B-meson decays generally have larger impact parameters to the beamline than background tracks, the reduced efficiency as a function of the impact parameter to the beamline results in a decreasing efficiency for reconstructing a B-meson decay as a function of its decay-time. This decreasing acceptance at higher decay-times is assumed to only affect Γ_s, a correction is intro-

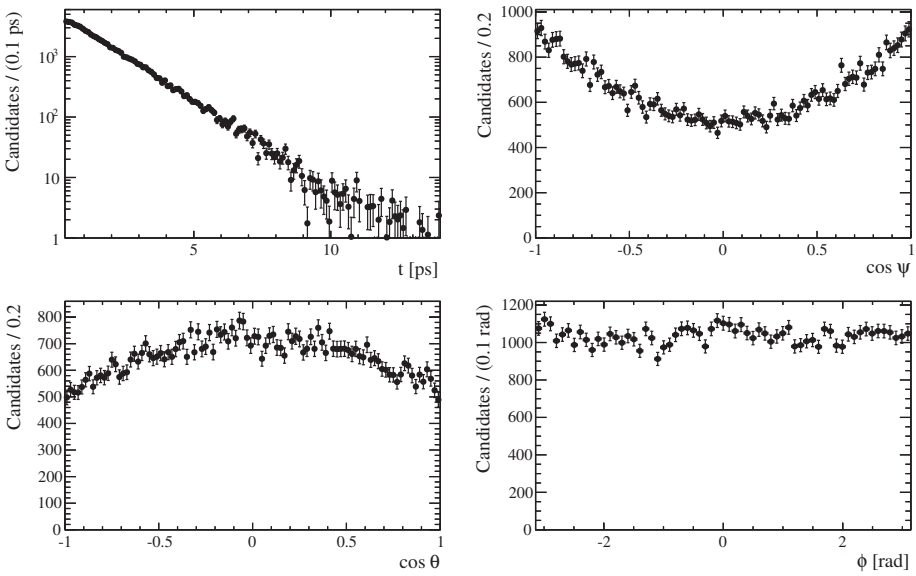


Figure 4.7: Background subtracted distributions of decay-time and decay-angles.

duced in the final fit. The correction is obtained from a separate fit, where events are by reweighing to correct for the decreasing acceptance at higher decay-times. The weights are extracted from data and the procedure is validated using simulated samples; it is described in more detail in section 5.1.

As described in section 4.1.6, the single track selections in HLT1 and the detached dimuon selection in HLT2 include requirements on observables which are correlated with a candidate's decay-time. At both HLT1 and HLT2 levels dedicated selections are available which have decay-time independent acceptances. By comparing the number of candidates selected by selections with decay-time dependent and decay-time independent acceptances, the relative acceptance is extracted from data in bins of the decay-time.

The decay-time dependent acceptance is taken into account in the likelihood fit by multiplying the time-dependent terms of the PDF with the decay-time acceptance according to:

$$\mathrm{PDF}_{\mathrm{acc}} \propto \varepsilon(t)\mathrm{PDF}_{\mathrm{phys}}, \tag{4.10}$$

where $\varepsilon(t)$ is the acceptance function; in this case a histogram. To propagate the statistical uncertainties of the acceptance per bin to the physics observables, the

determination of the acceptance is included in the likelihood fit. The details of this procedure are discussed in section 5.2.

4.5 Decay-Time Resolution

In general the decay-time resolution is defined as the width of the distribution of $t - t_{\text{true}}$, where t is the reconstructed decay-time, t_{true} is the true decay-time where the distribution is assumed to be Gaussian. The following sections describe how the decay-time resolution is modelled, how it affects time-dependent CP-violating observables and how per-candidate differences are taken into account.

4.5.1 Sources of Decay-Time Uncertainty

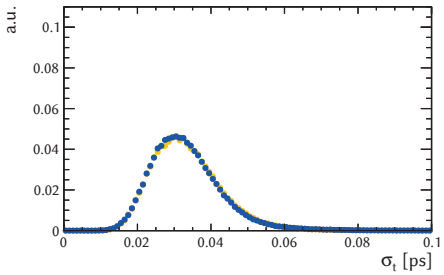
To calculate the contribution of the three-dimensional uncertainties of the positions of the primary vertex, the decay vertex and the $\mu^+\mu^-$ vertex to the estimated decay-time uncertainty of a candidate, the uncertainties of the vertices are assumed to be uncorrelated. Propagation of uncertainties is then applied to eq. (4.1) to give:

$$\sigma_{\text{vx}}^2 = \frac{M^2}{|p|^4}\mathbf{x}^T\mathbf{C}_{\text{vx}}\mathbf{x}, \tag{4.11}$$

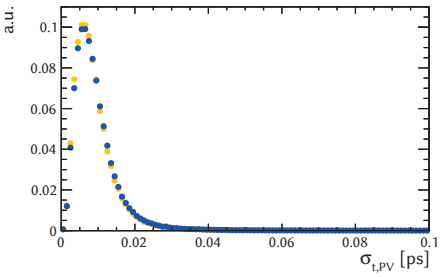
where $\mathbf{x}$ is the flight-distance vector of a candidate, and $\mathbf{C}_{\text{vx}}$ is the covariance matrix of the vertex whose contribution to the estimated decay-time uncertainty of the candidate is being calculated.

Since the primary vertex of a $b\bar{b}$-event consists on average of approximately 60 particle tracks [74, p. 179], in both the upstream and downstream directions, its position is better defined than the decay-vertex of a B-decay candidate. Its contribution to a B-decay candidate's decay-time is therefore expected to be small, which can be seen by comparing figs. 4.8a and 4.8b.

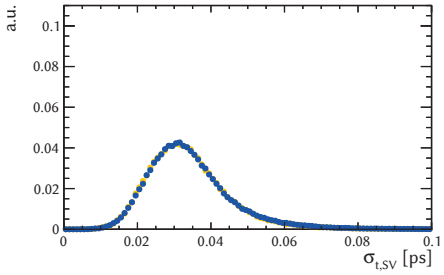
The secondary vertex consists of four tracks. As a result, its resolution is expected to be worse than that of the primary vertex and therefore the dominant contribution to the decay-time resolution. In addition, a significant variation is expected due to differences in the precision of reconstructed tracks from candidate to candidate. The contribution of the secondary vertex' uncertainty to the decay-time resolution is shown in fig. 4.8c. The fact that fig. 4.8a and fig. 4.8c are very similar supports the expectation that the secondary vertex' uncertainty is the dominant contribution to the decay-time resolution.



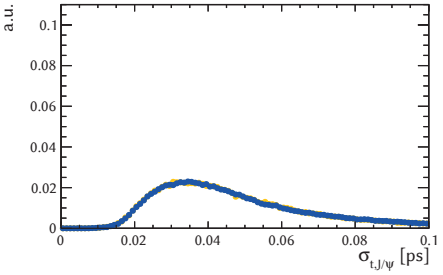
(a) Distribution of estimated decay-time uncertainty, σ_t.



(b) Contribution to σ_t of the primary vertex' uncertainty.



(c) Contribution to σ_t of the secondary vertex' uncertainty.



(d) Contribution to σ_t of the $\mu^+\mu^-$ vertex' uncertainty, the K^+K^- vertex is ignored.

Figure 4.8: Distribution of σ_t and contribution of uncertainties of several vertices to σ_t, for the 2011 (yellow) and 2012 (blue) samples.

The geometry of the tracks at the secondary vertex plays a role; tracks with a large opening angle generally form a vertex with a smaller uncertainty than parallel tracks. Since the φ resonance is just above the KK threshold, kaons from a $\varphi \to K^+K^-$ decay have little transverse momentum with respect to the momentum of the ϕ. Only the direction of the φ is therefore precisely measured, while the position of its vertex along the direction of its momentum remains largely unknown. The J/ψ resonance, however, is far above the $\mu^+\mu^-$ threshold, and the muons have nearly 2 GeV/c transverse momentum. As a result of the larger transverse momentum of the muons than of the kaons, the J/ψ vertex is expected to be measured more precisely than the K^+K^- vertex. If the position of the secondary vertex is replaced with the position of the $\mu^+\mu^-$ vertex in eq. (4.1), its contribution to the decay-time uncertainty is shown in fig. 4.8d. Comparing figs. 4.8c and 4.8d, it can be seen that the start of the turn-on of the distribution is the same in both figures, indicating that the minimal decay-time resolution is dominated by the $\mu^+\mu^-$ vertex' uncertainty.

The significant difference in the tails of these distributions shows that the inclusion of the K^+K^- vertex reduces the uncertainty of the vertex for those cases where, due to kinematics, the J/ψ vertex resolution is worse than usual.

If **x** and **p** are assumed to be parallel, eq. (4.1) simplifies to eq. (4.2). Propagation of uncertainties on d and p then gives:

$$\sigma_t^2 = \left(\frac{M}{p}\right)^2 \left(\sigma_d^2 + d^2 \left(\frac{\sigma_p}{p}\right)^2\right). \tag{4.12}$$

For single tracks, $\sigma_p/p \sim 0.55\%$ [74]. Assuming the four tracks share no hits, and their uncertainties are therefore uncorrelated, the B momentum uncertainty can be estimated to be $\sim 1.1\%$. From this it can be seen that the contribution to σ_t^2 of σ_p/p relative to that of σ_d scales with d, and that σ_t is, therefore, dominated by the contribution of the secondary vertex' resolution.

4.5.2 Decay-Time Resolution Model

The decay-time resolution model is assumed to consist of two Gaussian components whose widths are a function of the per-event estimated decay-time uncertainty.

$$\mathcal{R}(t|\sigma_t) \propto (1-f_2)\mathrm{e}^{-\frac{1}{2}\left(\frac{(t-\mu)}{\sigma_1}\right)^2} + f_2\mathrm{e}^{-\frac{1}{2}\left(\frac{(t-\mu)}{\sigma_2}\right)^2}, \tag{4.13}$$

It is taken into account in the likelihood fit by convoluting all decay-time dependent terms of the physics PDF with the decay-time resolution model. The

PDF then becomes:

$$\mathrm{PDF}(t, \boldsymbol{\Omega}|\sigma_t) = \mathcal{R}(t|\sigma_t) \otimes \mathrm{PDF}_{\mathrm{phys}}(t, \boldsymbol{\Omega}), \tag{4.14}$$

where $\mathrm{PDF}_{\mathrm{phys}}$ is given by eq. (1.24).

By propagating the uncertainties on all the parameters of eq. (4.1), the kinematic fit provides an estimate of the decay-time uncertainty σ_t and using this so-called per-event estimated decay-time uncertainty in the fit improves its performance by effectively giving greater weight to events with a smaller uncertainty. σ_t is incorporated into the resolution model of eq. (4.13) by replacing the widths of the Gaussian components, σ_1 and σ_2 by functions of σ_t: $\sigma_1(\sigma_t)$ and $\sigma_2(\sigma_t)$.

4.5.3 Calibration

The purpose of the calibration of the decay-time resolution model is to extract from data the values of it's parameters. For candidates with short "true" decay-times, the presence of a finite decay-time resolution sometimes results in negative reconstructed decay-times, which in turn results in an observed decay-time distribution which extends to $t < 0$. Since the decay-time resolution is the only source of such a negative time component, it can be used to calibrate the resolution model. The final selection contains a cut on the decay-time, so the prescaled selection is used instead to select the calibration sample. This sample, however, contains too few B decay candidates to allow calibration using the $\mathrm{B_s} \to \mathrm{J/\psi\,\varphi}$ decay only, such that an alternative sample of prompt-background candidates, that consist of two muons from a $\mathrm{J/\psi}$ decay and two random kaons, is required. This choice of calibration sample is motivated by the fact that about half of the selected candidates contain a $\mathrm{J/\psi} \to \mu^+\mu^-$ decay and the $\mathrm{J/\psi}$ vertex predominantly determines the estimated decay-time uncertainty. The calibration of the decay-time resolution model is described in detail in chapter 6.

4.5.4 Effect of Resolution on CP-Violating Observables

As shown in ref [75], the convolution of the decay-time distribution with the decay-time resolution model results in the dampening of the oscillations caused by the presence of time-dependent CP-violating observables, like ϕ_s. It has also been shown [75] that this dampening can be characterised by the so-called dilution, which is the real part of the Fourier transform of the decay-time resolution model. For a resolution model which consists of two Gaussians, this results in:

$$D = (1 - f_2)\mathrm{e}^{-\sigma_1^2 \Delta m_s^2/2} + f_2 \mathrm{e}^{-\sigma_2^2 \Delta m_s^2/2} \tag{4.15}$$

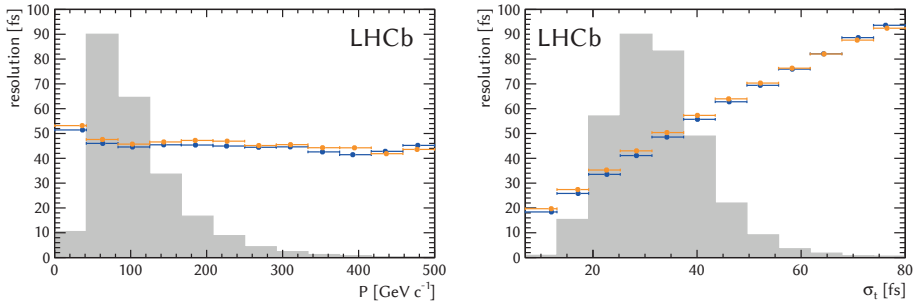


Figure 4.9: Effective decay-time resolution of a background subtracted sample of $B_s \to J/\psi\varphi$ decays as a function of B momentum (left) and estimated decay-time resolution (right).

Where D is the dilution of an event, σ_1 and σ_2 depend on σ_t, and Δm_s is the B_s oscillation frequency.

When events are weighted, the dilution of the total sample is equal to [76]:

$$D = \frac{1}{S_w}\sum_{i=0}^{N} w_i D_i, \tag{4.16}$$

where w_i is the event weight and S_w is the sum of these weights. Event weights could for example be sWeights or efficiency weights.

If per-candidate estimated decay-time uncertainties are introduced, it can be shown by minimising the ΔLog-Likelihood that the uncertainty on ϕ_s due to the presence of the decay-time resolution scales with the effective dilution, D_{eff}, where D_{eff} is defined as [76]:

$$D_{\text{eff}} = \sqrt{\frac{1}{S_w}\sum_{i=0}^{N} w_i D_i^2}, \tag{4.17}$$

To give an idea of the effect of the decay-time resolution on the observable amplitude of B_s-meson observations, an oscillation period may be compared to the decay-time resolution. Given that $\Delta m_s = 17.768 \pm 0.023 \pm 0.006\text{ps}^{-1}$ [77], $T = 2\pi/\Delta m_s = 353.62\text{fs}$. Using the model of eq. (4.13) with parameter values given in section 6.2.3 of chapter 6 together with the distribution of σ_t shown in fig. 4.8a, the effective dilution of the selected sample of $B_s \to J/\psi\varphi$ decays is calculated to be 0.7059 ± 0.0014. The width of a single Gaussian with equivalent dilution is approximately 47 fs, which is sufficiently narrower than a single oscillation period to allow the measurement of flavour oscillations.

Alternatively, the total dilution can be obtained numerically from the decay-time residual distribution using a discreet Fourier transform method, which calculates:

$$D_{\mathrm{FT}} = \frac{1}{S_w} \sum_{i=0}^{N} w_i \cos\left(\Delta m_s \delta_t\right), \tag{4.18}$$

where $\delta_t = t_{\mathrm{rec}} - t_{\mathrm{true}}$ is the decay-time residual.

Since the decay-time residual distribution is the only input to this method, it is model independent. It can therefore be used to asses the validity of a given model, by comparing the analytically calculated dilution to the one obtained directly from the decay-time residual distribution. To validate the numerical calculation of the dilution, a set of fast simulations have been performed, as described in appendix B. These show that $(D_{\mathrm{FT}} - D)/\sigma_{\mathrm{FT}} = -0.025 \pm 0.006$, which indicates that both methods give the same result with negligible bias. The validation procedure is described in more detail in appendix B.

To test the validity of the chosen decay-time resolution model, the dilution, D, is calculated of the sample of prompt-background candidates used for calibration of the resolution model. Applying eq. (4.16) gives $D = 0.6614 \pm 0.0012$, while eq. (4.18) gives 0.6604 ± 0.0038. The difference between these two results is -0.001 ± 0.004, which indicates that the chosen resolution model, consisting of two Gaussian components including the per-event estimated decay-time uncertainty, is appropriate for the determination of ϕ_s.

4.6 Angular Acceptance

The angles between the kaons and muons in a $B_s \to J/\psi\varphi$ decay need to be measured to be able to statistically disentangle CP-odd and CP-even components. Due to the geometry of the LHCb spectrometer, particles passing through the detector at a greater angle to the beamline are somewhat less likely to be reconstructed. The angular acceptance is implemented by multiplying the PDF with the angular-acceptance function $\varepsilon(\cos\theta_{\mathrm{K}}, \cos\theta_{\mu}, \phi_h)$, following [24, Sec 6.2.3]:

$$\mathrm{PDF}_{\mathrm{acc}}(t, \cos\theta_{\mathrm{K}}, \cos\theta_{\mu}, \phi_h) \propto \varepsilon(\cos\theta_{\mathrm{K}}, \cos\theta_{\mu}, \phi_h) \times \mathrm{PDF}_{\mathrm{phys}}(t, \cos\theta_{\mathrm{K}}, \cos\theta_{\mu}, \phi_h). \tag{4.19}$$

The angular-acceptance function is expanded in a basis of associated Legendre polynomials $P_i(\cos\theta_{\mathrm{K}})$ and real-valued spherical harmonics $Y_{lm}(\cos\theta_{\mu}, \phi_h)$:

$$\varepsilon(\mathbf{\Omega}) = \sum_{i,l,m} c^i_{lm} P_i(\cos\theta_{\mathrm{K}}) Y_{lm}(\cos\theta_{\mu}, \phi_h), \tag{4.20}$$

where $\mathbf{\Omega} = (\cos\theta_K, \cos\theta_\mu, \phi_h)$. The coefficients c^i_{lm} are calculated from simulated events using Monte Carlo integration.

A simulated sample of $B_s \to J/\psi\varphi$ decays is generated using the theoretical decay distribution $g(\mathbf{\Omega})$. Using the technique of Monte Carlo integration, an integral can be rewritten as a finite sum over a sample of generated events. The following then holds for an an integral over the product of the efficiency $\varepsilon(\mathbf{\Omega})$ and a general function $h(\mathbf{\Omega})$:

$$\int d\mathbf{\Omega}\varepsilon(\mathbf{\Omega})h(\mathbf{\Omega})g(\mathbf{\Omega}) \simeq \frac{1}{N_{\text{gen}}} \sum_{e\in\{\text{generated}\}} \varepsilon(\mathbf{\Omega}_e)h(\mathbf{\Omega}_e) \tag{4.21}$$

As each event is either accepted or not accepted, the per-event acceptance function $\varepsilon(\mathbf{\Omega}_e)$ is either 0 or 1, which allows the sum over generated events to be rewritten as a sum over accepted events:

$$\int d\mathbf{\Omega}\varepsilon(\mathbf{\Omega})h(\mathbf{\Omega})g(\mathbf{\Omega}) = \frac{1}{N_{\text{gen}}} \sum_{e\in\{\text{accepted}\}} h(\mathbf{\Omega}_e). \tag{4.22}$$

To obtain the parameters c^i_{lm} of eq. (4.20), the orthogonality relations of the chosen basis functions $P_i(\cos\theta_K)$ and $Y_{lm}(\cos\theta_\mu, \phi_h)$ are exploited. They included here for clarity:

$$\int_{-1}^{1} P^m_k P^m_l \, d\cos\theta_K = \frac{2(l+m)!}{(2l+1)(l-m)!}\delta_{kl} \tag{4.23}$$

$$\int_{-1}^{1} d\cos\theta_K \int_0^{2\pi} d\cos\theta_\mu d\phi_h Y_{lm}(\cos\theta_\mu, \phi_h) Y^*_{l'm'}(\cos\theta_\mu, \phi_h) = \delta_{ll'}\delta_{mm'}. \tag{4.24}$$

Choosing $h(\mathbf{\Omega})$ to be:

$$h(\mathbf{\Omega}) = \frac{2i+1}{2}\frac{P_i(\cos\theta_K)Y_{lm}(\cos\theta_\mu, \phi_h)}{g(\mathbf{\Omega})}, \tag{4.25}$$

and substituting eq. (4.20) in the left-hand side of eq. (4.22) leads to:

$$\int d\mathbf{\Omega} c^n_{jk} P_n(\cos\theta_K) Y_{jk}(\cos\theta_\mu, \phi_h) g(\mathbf{\Omega}) \times \left[\frac{2i+1}{2}\frac{P_i(\cos\theta_K)Y_{lm}(\cos\theta_\mu, \phi_h)}{g(\mathbf{\Omega})}\right] = c^i_{lm}, \tag{4.26}$$

and hence to:

$$c^i_{lm} = \frac{1}{N_{\text{gen}}} \sum_{e\in\{\text{accepted}\}} \frac{2i+1}{2}\frac{P_i(\cos\theta_K)Y_{lm}(\cos\theta_\mu, \phi_h)}{g(\mathbf{\Omega})}. \tag{4.27}$$

The coefficients of the expansion of the angular-acceptance function $\varepsilon(\mathbf{\Omega})$ can, therefore, be obtained from a finite sum of the function $h(\mathbf{\Omega})$ over accepted simulated events, as defined in eq. (4.25). Once the coefficients c^i_{lm} have been determined, the theoretical decay distributions for signal events are multiplied with $\varepsilon(\mathbf{\Omega})$ to correct for angular acceptance effects.

The angular acceptance can also be corrected for by using so-called normalisation weights for each of the angular functions $f_i(\mathbf{\Omega})$. These weights are obtained from eq. (4.22) and subsequently used to normalise the PDF [25]. The use of normalisation weights has been shown to be equivalent to the use of the acceptance function $\varepsilon(\mathbf{\Omega})$ in [24, App. F].

4.6.1 Iterative Reweighing

To take differences between the simulated sample and data into account, the simulated sample is reweighed to ensure the distributions of the momentum of the B_s, p_B, the K^+K^- invariant mass, M_{KK}, and kaon momenta, p_{K^+} and p_{K^-}, match those in data. The physics PDF used to generate simulated events is not necessarily correct, the non-resonant component in the K^+K^- invariant mass is, for example, not simulated, which could cause observable distributions to differ between the simulated and and data. This is taken into account by first reweighing the simulated sample such that the distribution of p_B and M_{KK} match those in background subtracted data. An iterative procedure is then employed, that proceeds as follows:

1. calculate normalisation weights,
3. perform the final fit to data to obtain parameter estimates,
4. reweigh the simulated sample, with event weights defined by $\varepsilon = \mathrm{PDF}(\mathbf{\Omega})/\mathrm{PDF}_{\mathrm{gen}}(\mathbf{\Omega})$, where the subscript "gen" indicates the PDF used to generate the simulated events,
5. reweigh simulated events such that the distributions of p_{K^+} and p_{K^-} match those in background subtracted data,
6. stop if $\Delta p/\sigma_p < 0.01$ for all physics parameters, p, continue with step 1., otherwise.

4.7 Flavour Tagging

For $B_s \to J/\psi\varphi$ decay, the oscillations that are observable in the time-dependent decay-rate in the presence of CP-violation, i.e. $\phi_s \neq 0$, have opposite sign for the different initial states of B_s and $\overline{B}_s$. The flavour of the initial state therefore needs

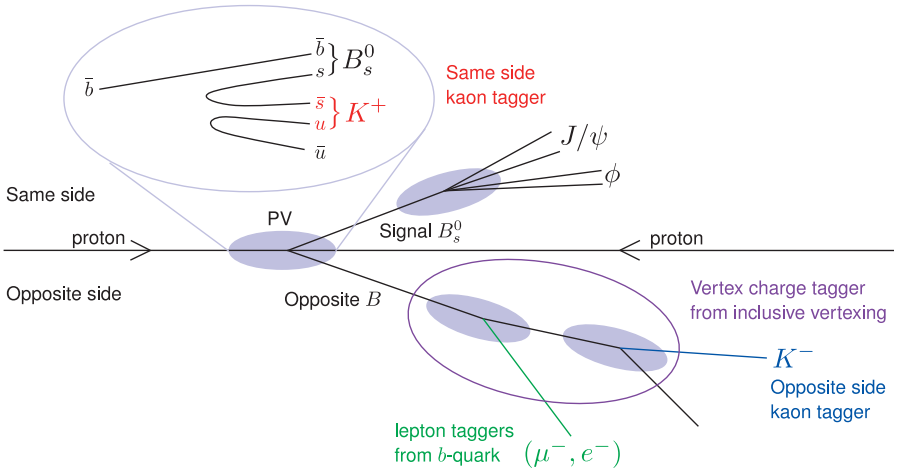


Figure 4.10: Schematic overview of the flavour tagging algorithms.

to be measured, which is accomplished by so-called flavour tagging algorithms. Two general methods of flavour tagging are used: opposite-side (OS) tagging and same-side (SS) tagging. Figure 4.10 shows a schematic overview of the tagging methods. Each tagging method results in a $\mathrm{B_s} \to \mathrm{J/\psi\,\varphi}$ candidate being tagged as originally being a $\mathrm{B_s}$, a $\overline{\mathrm{B}}_\mathrm{s}$, or remaining untagged, with an associated tagging decision $q = \{1, -1, 0\}$, respectively.

The probability that an algorithm assigns a non-zero flavour tag to a candidate is represented by that algorithm's efficiency, $\varepsilon_{\mathrm{tag}}$, which is defined as:

$$\varepsilon_{\mathrm{tag}} = \frac{R + W}{R + W + U}, \tag{4.28}$$

where R, W and U are the number of correctly tagged, wrongly tagged and untagged candidates, respectively.

The probability that a $\mathrm{B_s} \to \mathrm{J/\psi\,\varphi}$ candidate has been tagged with the wrong flavour, η, is estimated event-by-event by the algorithms, using a neural network trained on part of a data sample of background subtracted $\mathrm{B^+} \to \mathrm{J/\psi K^+}$ decays. To obtain the calibrated wrong-tag probability, ω, from the estimated wrong-tag probability, η, the following relations are defined:

$$\omega = p_0 + p_1(\eta - \bar{\eta}) \quad \text{for an initial } \mathrm{B_s} \text{ candidate,} \tag{4.29}$$

$$\bar{\omega} = p_0 - p_1(\eta - \bar{\eta}) \quad \text{for an initial } \overline{\mathrm{B}}_\mathrm{s} \text{ candidate,} \tag{4.30}$$

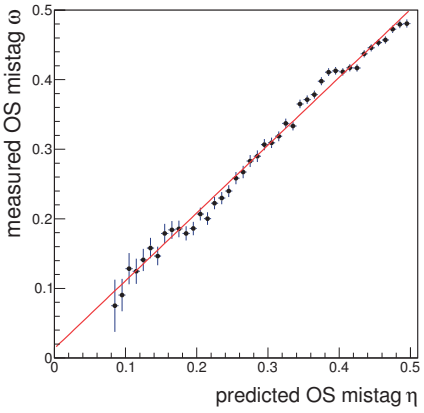


Figure 4.11: Measured (ω) versus estimated (η) wrong-tag probability obtained from a sample of $B^+ \to J/\psi K^+$ decays. The line represents a linear parameterisation of the form of eq. (4.31), obtained from a fit to the values of ω. Figure provided by [78].

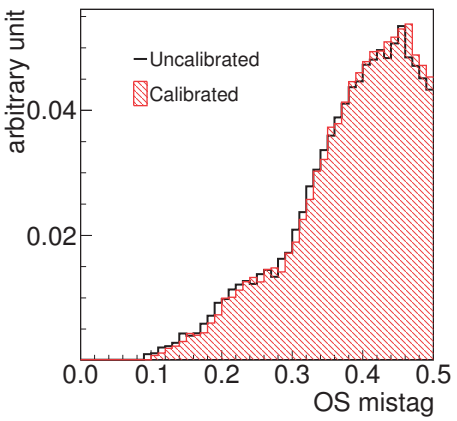


Figure 4.12: Distribution of η of the combined opposite-side tagging algorithms for $B_s \to J/\psi\varphi$ signal candidates before and after calibration.

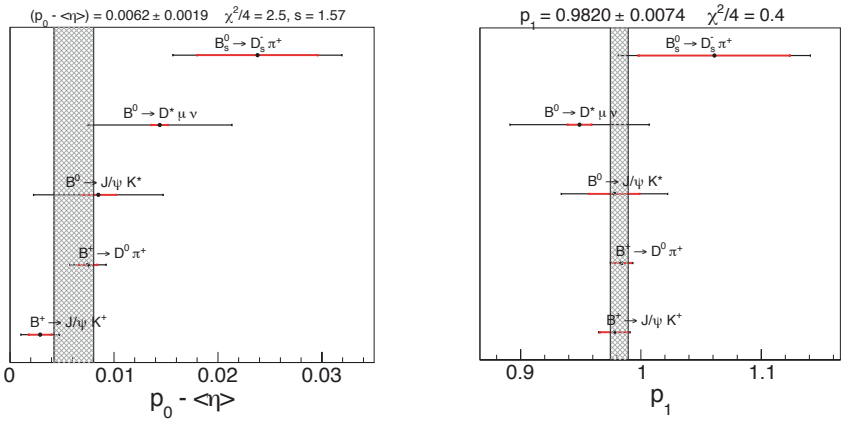
where $\bar{\eta}$ is the average estimated wrong-tag probability, and p_0 and p_1 are defined as:

$$p_0 = \frac{p_{0,b} + p_{0,\bar{b}}}{2}$$
$$p_1 = \frac{p_{1,b} + p_{1,\bar{b}}}{2} \tag{4.31}$$

where the parameters with a b subscript refer to initial B_s candidates and those with a $\bar{b}$ subscript refer to initial $\overline{B}_s$ candidates. p_0 and p_1 are, thus, averaged between initial B_s and $\overline{B}_s$ candidates. Asymmetries in the parameters between samples of initial B_s and initial $\overline{B}_s$ candidates were found to be small and are neglected [69, Sec. 8].

4.7.1 Opposite-Side Taggers

The opposite-side tagging relies on the assumption that b and $\bar{b}$ quarks, which are produced as a pair, hadronise independently. After decay of the meson or



(a) parameter: $p_0 - \langle\eta\rangle$

(b) parameter: p_1

Figure 4.13: Parameters of the relations listed in eq. (4.31) obtained from control channels, and their average. Figures provided by [78].

baryon that the non-signal b or $\overline{\text{b}}$ quark hadronises into, the charge of — some of — the decay products is correlated to the initial flavour of the b or $\overline{\text{b}}$ quark, and, therefore, to the initial flavour of the signal B_s meson. The correlation is lost if the non-signal b or $\overline{\text{b}}$ quark hadronises into a B_s meson, due to its rapid oscillations.

Separate algorithms exist which focus on either muons and electrons from semileptonic decays, charged kaons from $b \to c \to s$ transitions or inclusive charge of the decay-vertices. The decisions and wrong-tag probabilities provided by each of these algorithms are combined event-by-event in a neural network to give the final opposite-side tagging decision and estimated wrong-tag probability.

In addition to optimisation and calibration using a background subtracted sample of $\text{B}^+ \to \text{J}/\psi\text{K}^+$ decays, the wrong-tag probability, η, is calibrated using a background subtracted sample of $\text{B}^0 \to \text{J}/\psi\text{K}^*$ decays, and by measuring the flavour oscillations in a sample of $\text{B}^0 \to \text{D}^*\ell\nu$ decays. The calibration from the sample of $\text{B}^+ \to \text{J}/\psi\text{K}^+$ decays is used in the final fit, while the calibrations obtained from the other samples are used for cross-checks.

Calibration is performed with a number of control channels in addition to the part of the $\text{B}^+ \to \text{J}/\psi\text{K}^+$ sample not used for optimisation: samples of $\text{B}^+ \to \text{D}^0\pi^+$, $\text{B}^0 \to \text{J}/\psi\text{K}^*$, $\text{B}^0 \to \text{D}^*\ell\nu$ and $\text{B}_\text{s} \to \text{D}_\text{s}^-\pi^+$. For $\text{B}^+ \to \text{J}/\psi\text{K}^+$

and $B^+ \to D^0\pi^+$ decays, the charge of the final state hadrons directly determines whether a b or $\bar{b}$ quark was produced, and ω is obtained by counting the fraction of correctly tagged decays in bins of η. For each of the samples of neutral B mesons, a simultaneous fit to the decay-time distribution in bins of η is performed to obtain ω. The parameters of the relations given in eq. (4.31) are subsequently obtained from a fit to the measured ω as a function of η. The parameters obtained from all control samples are shown in fig. 4.13, and are combined using the PDG averaging method [14, Ch 1. Sec. 5.2.2] to give the final calibration parameters listed in table 4.6.

Figure 4.11 shows the calibration obtained from a sample of $B^+ \to J/\psi K^+$ candidates, and fig. 4.12 shows the distributions of the estimated and calibrated wrong-tag probabilities, η and ω. Figure 4.13 shows the parameters obtained from each control channel and their average; the average values are additionally listed in table 4.6.

parameter	OS	SS
$\langle\eta\rangle$	0.38	0.44
$p_0 - \langle\eta\rangle$	0.0062 ± 0.0044	0.005 ± 0.005
p_1	0.982 ± 0.035	0.976 ± 0.091

Table 4.6: Parameters of eq. (4.31) obtained from a fit to the per-event wrong-tag probabilities measured with control samples. Systematic uncertainties have been added to the statistical uncertainties in quadrature to obtain the reported uncertainties. Systematic uncertainties were obtained by taking the quadratic sum of observed differences between repeated calibrations with control samples reweighed to the kinematic properties of several signal decays.

4.7.2 Same-Side Taggers

The same-side tagging algorithm relies on the hadronisation of the partner of the spectator quark in B_s mesons: an s or $\bar{s}$ quark. If this partner quark hadronises into a charged kaon, the charge of the kaon is correlated with the flavour of the produced B_s meson. Additionally, the kaon containing the partner of the spectator quark is produced close to the B_s meson in phase space. The same-side tagging algorithm selects a sample of kaons that are close in phase-space to the B_s meson, and uses a neural network to select from this sample the kaon, that is most likely to have been created from the partner of the spectator quark. The

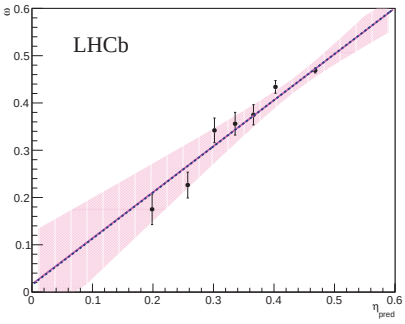


Figure 4.14: Measured (ω) versus estimated (η) wrong-tag probability obtained from a simultaneous fit in bins of η to the decay-time distribution of a sample of $B_s \to D_s^- \pi^+$ decays. The line represents a linear parameterisation of the form of eq. (4.31), obtained from a fit to the values of ω

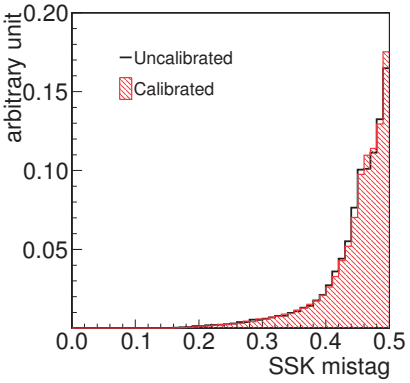


Figure 4.15: Distribution of η of the same-side tagging algorithm for $B_s \to J/\psi\varphi$ signal candidates before and after calibration.

charge of this kaon then identifies the flavour of the signal B_s meson at production.

The same-side tagging algorithms are optimised on simulated samples and calibrated by fitting the flavour oscillations in a sample of $B_s \to D_s^- \pi^+$ decays. Figure 4.14 shows the calibration obtained from a sample of $B_s \to D_s^- \pi^+$ candidates, and fig. 4.15 shows the distributions of the estimated and calibrated wrong-tag probability, η and ω, of the same-side tagging algorithms. The parameters obtained from a fit of the relations given in eq. (4.31) to distribution of the measured wrong-tag probabilities are listed in table 4.6.

4.7.3 Tagging Performance

To evaluate the performance of the flavour tagging algorithms, the effective tagging power, $\varepsilon(1-2\omega)^2$, is generally quoted in addition to the efficiency, ε_{tag} and wrong-tag probability, ω of a tagging algorithm. It gives the fraction of the signal sample that is statistically equivalent to a sample tagged with $\omega = 0$. Table 4.7 lists the performance of the tagging algorithms for the background subtracted sample of $B_s \to J/\psi\varphi$ candidates. The $B_s \to J/\psi\varphi$ sample is split into four in-

category	fraction (%)	ε (%)	$\varepsilon(1-2\omega)^2$ (%)
OS-only	20.5	13.74 ± 0.11	1.19 ± 0.06
SS-only	54.0	36.02 ± 0.15	0.84 ± 0.11
OS$\wedge$SS	25.5	17.00 ± 0.12	1.70 ± 0.08
Total	100	66.76 ± 0.22	3.73 ± 0.15

Table 4.7: Tagging performances for the exclusively OS, exclusively SS, and doubly tagged events, for $B_s \to J/\psi\varphi$ signal candidates, after calibration of the wrong-tag probability. The column "fraction" reports the fraction of events in each sample out of the entire tagged sample. The errors on $\varepsilon(1-2\omega)^2$ take the uncertainties on the calibration parameters into account.

dependent samples according to the tagging decisions of the opposite-side and same-side algorithms: exclusively OS tagged ($q^{OS} \neq 0 \wedge q^{SS} = 0$), exclusively SS tagged ($q^{OS} = 0 \wedge q^{SS} \neq 0$), doubly tagged ($q^{OS} \neq 0 \wedge q^{SS} \neq 0$) and, finally, untagged ($q^{OS} = 0 \wedge q^{SS} = 0$) events, which comprise approximately 33% of the sample. The performance of the tagging for $B_s \to J/\psi\varphi$ signal candidates is listed in table 4.7. The total tagging efficiency and tagging power are $(66.76 \pm 0.22)\%$ and $(3.73 \pm 0.15)\%$, respectively.

4.7.4 Incorporating Tagging in the Signal PDF

To incorporate the tagging in the time and decay-angle dependent PDF, the following relations are used:

$$\begin{aligned}\mathrm{PDF}(t, \boldsymbol{\Omega}|q^{OS}, q^{SS}, \eta^{OS}, \eta^{SS}) \propto & (1 + q^{OS}(1 - 2\omega^{OS})) \\ & \times (1 + q^{SS}(1 - 2\omega^{SS}))\mathrm{PDF}(t, \boldsymbol{\Omega}|B_s) \\ & + (1 - q^{OS}(1 - 2\bar{\omega}^{OS})) \\ & \times (1 - q^{SS}(1 - 2\bar{\omega}^{SS}))\mathrm{PDF}(t, \boldsymbol{\Omega}|\overline{B}_s), \end{aligned} \tag{4.32}$$

where ω and $\bar{\omega}$ are derived from eq. (4.31), and $\mathrm{PDF}(t, \boldsymbol{\Omega}|B_s)$ and $\mathrm{PDF}(t, \boldsymbol{\Omega}|\overline{B}_s)$ are the decay-time and decay-angle dependent decay rates for initial B_s and $\overline{B}_s$ particles obtained by setting q_T in eq. (1.24) to 1 or -1, respectively. By explicitly considering $q^{OS} = 0$ and $q^{SS} = 0$ as tagging decisions, eq. (4.32) describes the decay rates for both tagged and untagged samples.

The correlation between OS and SS tagging decisions and wrong-tag probabilities were found to be negligible, as were tagging efficiency and wrong-tag probability asymmetries between candidates tagged as B_s and candidates tagged as

$\overline{\mathrm{B}}_\mathrm{s}$. In addition, $\mathrm{PDF}(t, \boldsymbol{\Omega}|\mathrm{B}_\mathrm{s})$ and $\mathrm{PDF}(t, \boldsymbol{\Omega}|\overline{\mathrm{B}}_\mathrm{s})$ are normalised separately, which ensures that a possible $\mathrm{B}_\mathrm{s}/\overline{\mathrm{B}}_\mathrm{s}$ production asymmetry has no effect on the physics parameters of interest.

4.8 Observable PDF

To create the experimentally observable PDF for the background subtracted sample of $\mathrm{B}_\mathrm{s} \to \mathrm{J}/\psi\varphi$ decays, the signal PDF given by eq. (1.24) is augmented to take four effects into account: decay-time resolution, decay-time acceptance, angular acceptance and flavour tagging. The angular resolution is ignored and a corresponding systematic uncertainty is included in the final result, as described in section 7.2.9.

To take the decay-time resolution into account, the signal PDF is convoluted with the resolution model given in eq. (4.13). The decay-time acceptance resulting from the HLT selections is subsequently taken into account by multiplying the PDF — including the decay-time resolution model — with the decay-time acceptance histogram described in section 4.4, and in more detail in section 5.2. The decay-time acceptance resulting from the lower efficiency of the Velo track finding to reconstruct tracks with larger impact parameters with respect to the beamline is taken into account by reweighing event appropriately, as described in section 4.4, and in more detail in section 5.1.

To take the angular acceptance into account, the decay-angle dependent functions of the PDF are multiplied with the angular acceptance according to eq. (4.19). To take the final experimental effect — the flavour tagging — into account, the PDF is augmented according to eq. (4.32), and the data sample is split in two parts, that are tagged as either initially being a B_s or $\overline{\mathrm{B}}_\mathrm{s}$, to allow a simultaneous fit.

The PDF describing the observable decay-rates for $\mathrm{B}_\mathrm{s} \to \mathrm{J}/\psi\varphi$ candidates either tagged — as B_s or $\overline{\mathrm{B}}_\mathrm{s}$ — or untagged becomes:

$$\begin{aligned}
&\mathrm{PDF}(t, \boldsymbol{\Omega}|q^{\mathrm{OS}}, q^{\mathrm{SS}}, \eta^{\mathrm{OS}}, \eta^{\mathrm{SS}}, \sigma_t) \propto \\
&\quad (1 + q^{\mathrm{OS}}(1 - 2\omega^{\mathrm{OS}}))(1 + q^{\mathrm{SS}}(1 - 2\omega^{\mathrm{SS}})) \\
&\quad \times \left\{\varepsilon(\boldsymbol{\Omega})\varepsilon(t)\left[\mathcal{R}(t|\sigma_t) \otimes \mathrm{PDF}(t, \boldsymbol{\Omega}|\mathrm{B}_\mathrm{s})\right]\right\} \\
&\quad + (1 - q^{\mathrm{OS}}(1 - 2\bar{\omega}^{\mathrm{OS}}))(1 - q^{\mathrm{SS}}(1 - 2\bar{\omega}^{\mathrm{SS}})) \\
&\quad \times \left\{\varepsilon(\boldsymbol{\Omega})\varepsilon(t)\left[\mathcal{R}(t|\sigma_t) \otimes \mathrm{PDF}(t, \boldsymbol{\Omega}|\overline{\mathrm{B}}_\mathrm{s})\right]\right\},
\end{aligned} \tag{4.33}$$

where $\varepsilon(t)$ is the decay-time acceptance function; $\varepsilon(\boldsymbol{\Omega}))$ is the angular acceptance function defined in eq. (4.20); $\mathcal{R}(t|\sigma_t)$ is the decay-time resolution model

defined in eq. (4.13); ω and $\bar{\omega}$ are derived from eq. (4.31); and PDF$(t, \mathbf{\Omega}|\mathrm{B_s})$ and PDF$(t, \mathbf{\Omega}|\overline{\mathrm{B}}_\mathrm{s})$ are obtained by setting q_T in eq. (1.24) to 1 or -1, respectively.

The origins of the decay-time dependent efficiency, $\varepsilon(t)$, and how it is taken into account in the PDF are discussed in more detail in the following chapter. Chapter 6 then describes the calibration of the decay-time resolution model $\mathcal{R}(t|\sigma_t)$, and the results of the analysis and the systematic uncertainties on the parameter estimates are given and discussed in chapter 7.

Chapter 5

Decay-Time Acceptance

In LHCb, many analyses select candidates using requirements on decay-time correlated quantities. If a decay-time dependent likelihood fit to the sample is to be performed, the decay-time dependent efficiency induced by the selection requirements needs to be known and modelled. In LHCb, this decay-time dependent efficiency function is referred to as the decay-time acceptance.

Three sources of a decay-time dependent acceptance have been observed so far. The first is a result of the selection requirement of $t > 0.3$ ps applied as part of the final selection of $B_s \to J/\psi\,\varphi$ candidates. Since the lower bound of the range in which the final fit is performed was chosen to be equal to the value of the requirement, the offline selection does not result in a decay-time dependent acceptance.

For all data samples, an implicit source of a decay-time dependent acceptance is the result of a decreased efficiency of the Velo track finding algorithms for tracks with larger distances to the beamline, which results in a decay-time dependent acceptance at large decay-times. Since this effect is present for all reconstructed tracks, a relative acceptance does not account for it and it is therefore determined separately, and taken into account by correcting the lifetime in the fit. This procedure is described in section 5.1.

The selection requirement on the decay-length significance, eq. (3.11), of the J/ψ applied as part of the HLT2 detached dimuon selection is correlated with the B_s decay-time and results in a non-trivial acceptance function for the entire sample. The impact parameter and impact parameter significance selection requirements which are part of the single track HLT1 selections, affect only the subsample of candidates that have been selected by these HLT1 selections and not by the HLT1 high-mass dimuon selection. Since a so-called unbiased sample,

which has been selected without requirements on decay-time correlated quantities, is available, a relative acceptance can be extracted and included in the likelihood. This procedure is described in section 5.2.

If an unbiased sample is not available, the acceptance function can be extracted from simulated samples, which may lead to systematical uncertainties due to the presence of differences between MC and data. Alternatively the acceptance can be extracted from data for each candidate and as such taken into account in the likelihood. This alternative procedure is referred to as swimming and described in section 5.3.

5.1 Decay-Time Acceptance from Reconstruction

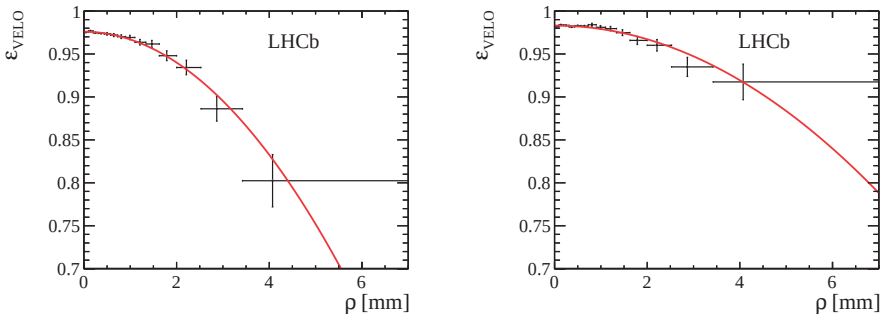
For a given track, its distance of closest approach (DOCA) with respect to the beamline, is obtained from the standard formula for the distance between two skew lines and given by:

$$\mathrm{DOCA}_z \equiv \frac{|(\mathbf{d}-\mathbf{v})\cdot(\mathbf{p}\times\hat{\mathbf{z}})|}{|\mathbf{p}\times\hat{\mathbf{z}}|}, \tag{5.1}$$

where $\mathbf{p}$ is the momentum vector of a track with a measurement at $\mathbf{d}$; $\hat{\mathbf{z}}$ is a unit vector along the beamline, which is parallel to the z-axis of the global LHCb coordinate system and $\mathbf{v}$ is the origin of the Velo coordinate system. Once the LHC has brought the beams into collision and declared stable conditions, the Velo (and it's local coordinate system) are centred on the interaction point.

Offline, the Velo track finding algorithm proceeds in three distinct steps and in each step only hits not previously assigned to tracks are considered. In the first step, track seeds that have hits in four consecutive R-sensors are used. Seeds are then extended and hits in Φ-sensors are searched for and added. If sufficient R and Φ hits are found, and the resulting track is of sufficient quality, it is accepted. The second step starts from track seeds that have hits in three consecutive R-sensors; it then proceeds in the same manner as the first step. The third step starts from seeds consisting of hits in the Φ-sensors and extends them before adding hits in R-sensors.

The first two steps assume that tracks originate from the region around the beamline, while the third step does not. The assumption is that, given two track candidates, the candidate with the smallest DOCA_z is the better candidate [74, 54]. This results in a decreased efficiency for finding a track as a function of its DOCA_z, which is not entirely compensated by the third step. Due to the non-zero lifetime and the large boost of B-mesons, tracks from most B-decays generally do



	a	c$[\,mm^{-2}]$
online	0.9759 ± 0.0005	-0.0093 ± 0.0007
offline	0.9831 ± 0.0004	-0.0041 ± 0.0005

Figure 5.1: Velo-track finding efficiency for kaon tracks reconstructed using the online (top left) and offline (top right) algorithms as a function of the kaon DOCA$_z$, as defined in eq. (5.1). The red solid lines show the result of an unbinned maximum likelihood fit (bottom) using the parameterisation given by eq. (5.2) to the background subtracted data (black points). Note that a typical DOCA$_z$ is 0.1–0.2 mm.

not originate from the beamline and therefore have a somewhat lower efficiency of being reconstructed [79]. Since a B-decays cannot be reconstructed as soon as one of it's daughters' tracks is not reconstructed, the reduction in efficiency at large decay-times increases with the number of charged final-state particles.

The Velo track finding algorithms are configured differently in HLT than in the final reconstruction. At the beginning of HLT1, only Velo track seeds which have hits in four consecutive sensors are considered. At the start of the event reconstruction performed in HLT2, an additional search for Velo tracks is performed which accepts track seeds with hits in three consecutive sensors and also searches for tracks without the constraint that tracks should originate from the beamline. This additional step brings the Velo reconstruction used in HLT2 on par with the offline Velo reconstruction.

By using a sample of $B^+ \to J/\psi K^+$ decays, where Velo information was ignored during the reconstruction of the kaon track, the single track reconstruction efficiency can be determined as a function of DOCA$_z$ by testing how often the kaon track is found once the Velo information is included. The different configuration of the Velo track finding algorithms in HLT1 with respect to the offline

Correction	2011	2012
$\Delta\tau_{B_s}$ (fs)	14.8	23.2
$\Delta\tau_{B_s}^{MC}$ (fs)	5.0	4.2
$\Delta\tau_{B_s}^{\text{total}}$ (fs)	19.8	27.4
β ($\times 10^{-3}$ ps^{-1})	-9.0 ± 2.2	-12.4 ± 1.9

Table 5.1: Summary of corrections due to the reconstruction induced decay time acceptance, reproduced from [80].

reconstruction is taken into account by determining the efficiency separately for the muon tracks, which are used in HLT1, and for the offline reconstructed kaon tracks. The observed efficiencies for the online and offline configurations of the Velo track finding algorithm are shown in Figure 5.1, together with a parameterisation defined by:

$$\varepsilon_{Velo}(\text{DOCA}_z) = a(1 + c\,\text{DOCA}_z^2), \tag{5.2}$$

where a and c are determined from fits to the observed distributions; the result of these fits is given in fig. 5.1.

To take the per track efficiencies into account in the final fit, efficiency correction weights, w_{Velo}, are introduced, that are obtained from the single single-track reconstruction efficiency, ε_{Velo}, according to:

$$w_{\text{Velo}} = 1/\left(\varepsilon_{Velo,online}^{\mu^+}\,\varepsilon_{Velo,online}^{\mu^-}\,\varepsilon_{Velo,offline}^{K^+}\,\varepsilon_{Velo,offline}^{K^-}\right), \tag{5.3}$$

where the μ and K superscripts indicate online reconstructed muons and offline reconstructed kaons, respectively. The difference in lifetime observed between fits with and without reconstruction efficiency weights, β, is taken as a correction and is incorporated in the final fit by replacing the average lifetime Γ in the observable PDF by $\Gamma' = \Gamma - \beta$.

To validate the weighting procedure, it is performed on simulated candidates. The difference observed between the generated lifetime and the one recovered by including w_{Velo} is added to β. The contributions to the final beta factors and their values are listed in table 5.1.

Differences in the decay-time dependent acceptance between the HLT1 single track selections and the HLT1 high-mass dimuon selection are taken into account by the relative acceptance, which is described in the following section.

5.2 Fitting for the Decay-Time Acceptance

The HLT1 single track selections, as described in section 3.2.8, contain selection requirements on the impact parameter and impact parameter significance of the candidate track. If tracks originate from a $B_s \to J/\psi\varphi$ decay, these quantities are correlated with the candidate's decay time and their inclusion in the trigger selection results in a decay-time dependent acceptance. This is similarly the case for the selection requirement on the decay-length significance of $\mathrm{DLS}(J/\psi) > 3$ in the HLT2 detached dimuon selection described in section 3.3.2.

At both the HLT1 and HLT2 level, selections are available which have a decay-time independent acceptance. The high-mass dimuon selection performs this role in HLT1 while in HLT2 an additional dimuon selection is implemented which does not include the requirement of $\mathrm{DLS}(J/\psi) > 3$, but only runs on approximately 20% of the input events. Since the only difference between the two HLT2 selections is the decay-length significance cut, the number of candidates which have been accepted by a combination of these selections is used to determine the relative decay-time dependent acceptance in bins of decay-time. Given the number of candidates which have been accepted by the respective trigger selections in each bin the appropriate relations are:

$$\begin{aligned} N_{\mathrm{pd}} &= \varepsilon_{\mathrm{p}}^2 \varepsilon_{\mathrm{d}}^2 N, \\ N_{\mathrm{p}} &= (1-\varepsilon_{\mathrm{d}}^2)\varepsilon_{\mathrm{p}}^2 N, \\ N_{\mathrm{d}} &= (1-\varepsilon_{\mathrm{p}}^2)\varepsilon_{\mathrm{d}}^2 N, \end{aligned} \tag{5.4}$$

where bin indices have been omitted; N_{d} and N_{p} are the number of candidates accepted by the detached and prescaled HLT2 selections with and without the requirement of $\mathrm{DLS}(J/\psi) > 3$, respectively; N_{pd} is the number of candidates accepted by both selections; N is the number of candidates before selection and the superscript 2 refers to HLT2;

The HLT1 single track selections uses single tracks, while the high-mass dimuon selection requires a combination of two tracks. Additionally, the single track selections contain requirements which are different from the high-mass dimuon selection, but which do not result in a decay-time dependent acceptance; the total and transverse momentum of tracks are examples of such requirements. Due to these differences the number of candidates accepted by both selections cannot be used. Candidates are therefore divided into two mutually exclusive categories, candidates selected by the HLT1 dimuon high mass selection are labelled as unbiased, while candidates selected by the HLT1 single track selections and not by the HLT1 dimuon high mass selection are label-ed as exclusively

biased. The expressions for the number of candidates in each category in bins of decay-time respectively become:

$$\begin{aligned} N_{\mathrm{u}} &= \varepsilon_{\mathrm{u}}^{1} N, \\ N_{\mathrm{e}} &= \varepsilon_{\mathrm{e}}^{1} N, \end{aligned} \tag{5.5}$$

where bin indices have been omitted, the index u refers to candidates accepted by the HLT1 high mass dimuon selection, the index e to candidates accepted by at least one of the HLT1 single track selection and not by the HLT1 high mass dimuon selection and the superscript 1 refers to HLT1.

The decay-time dependent acceptance is taken into account in the likelihood fit by multiplying the decay-time dependent terms of the PDF by the acceptance bin-by-bin. For each bin an independent set of equations as defined in eqs. (5.4) and (5.5) applies. To allow (fast) analytical calculation of the required integrals, the acceptance is expressed as a histogram. Discrete observables have been added to the data sample, indicating by which trigger selections a candidate has been selected. The value of the acceptance in each bin in decay-time is calculated according to the value of these discrete observables. This implies that the PDF is conditional on the discrete observables and the PDF contains no information about the total number of candidates accepted in each bin. Combining the selection categories for HLT1 and HLT2 results in a total of six categories:

$$\begin{aligned} N_{\mathrm{u,p}} &= \varepsilon_{\mathrm{u}}^{1} \varepsilon_{\mathrm{p}}^{2} (1 - \varepsilon_{\mathrm{d}}^{2}) N, & N_{\mathrm{u,d}} &= \varepsilon_{\mathrm{u}}^{1} (1 - \varepsilon_{\mathrm{p}}^{2}) \varepsilon_{\mathrm{d}}^{2} N, \\ N_{\mathrm{u,pd}} &= \varepsilon_{\mathrm{u}}^{1} \varepsilon_{\mathrm{p}}^{2} \varepsilon_{\mathrm{d}}^{2} N, & N_{\mathrm{e,p}} &= \varepsilon_{\mathrm{e}}^{1} \varepsilon_{\mathrm{p}}^{2} (1 - \varepsilon_{\mathrm{d}}^{2}) N, \\ N_{\mathrm{e,d}} &= \varepsilon_{\mathrm{e}}^{1} (1 - \varepsilon_{\mathrm{p}}^{2}) \varepsilon_{\mathrm{d}}^{2} N, & N_{\mathrm{e,pd}} &= \varepsilon_{\mathrm{e}}^{1} \varepsilon_{\mathrm{p}}^{2} \varepsilon_{\mathrm{d}}^{2} N, \end{aligned} \tag{5.6}$$

where the superscripts 1 and 2 refer to HLT1 and HLT2, respectively and decay-time bin indices have been omitted.

To allow the determination of the decay-time dependent HLT acceptance as part of the likelihood fit, the information about the total number of candidates accepted by each HLT selection per bin in decay-time is reintroduced into the log-likelihood by adding a specifically designed additional term to it, which acts as a constraint on the bin-by-bin acceptance values. The number of candidates in the categories listed in eq. (5.6) are assumed to be distributed according to Poisson distributions. Since sWeights are used to obtain the observable distributions for signal candidates, they need to be taken into account when counting the number of candidates in each category. The contribution to the likelihood by a Poisson distribution is given by:

$$\log P(N) = N \log \hat{N} - \hat{N}, \tag{5.7}$$

where N is the number of observed candidates in a category and the $\hat{N}$ the estimated number of candidates in a category. To take the sWeights into account, the Poisson distribution is modified to have a mean equal to the sum of weights in a category and a variance equal to the sum of squared weights [81]. The logarithm of the modified distribution is given by:

$$\log P(N) = \frac{s_w}{s_{w^2}} \left(s_w \log \hat{N} - \hat{N} \right), \tag{5.8}$$

where s_w is the sum of weights and s_{w^2} is the sum of squared weights; bin and category indices have been omitted.

Since only candidates which pass the final selection are of interest, the acceptance is also defined relative to the final selection. The turn-on of the time dependent efficiency of the HLT2 detached dimuon selection is mostly located below the final selection requirement of $t > 0.3$ ps, which results in an average relative acceptance for the $\mathrm{B_s} \to \mathrm{J/\psi\,\varphi}$ sample close to 100%. As a result, N_p is rather small and to avoid problems, $N_\mathrm{u,p}$ and $N_\mathrm{e,p}$ are assumed to be the same, which leaves five categories. The final addition to the likelihood is the sum over all bins in decay-time and the remaining five categories.

By including the constraint given in eq. (5.8) for all bins and trigger categories in the likelihood of the final fit, the per-bin efficiencies of the HLT1 single track and the HLT2 detached dimuon acceptance histograms can be let free and determined by the fit. The values for the per-bin efficiencies and their uncertainties as determined by the final fit are shown in fig. 5.2. The strong turn-on of the HLT1 single track selections is clearly visible. The fluctuations visible in the HLT1 single track shapes are likely a consequence of number of bins used (40). The similarity of the curves for the 2011 and 2012 samples is a result of the fact that no changes were made to the trigger selections for the duration of LHC Run-I.

5.3 Per-Event Decay-Time Acceptance (Swimming)

If unbiased samples — selected without requirements on decay-time correlated quantities — are not available, or have insufficient statistics, an alternative method can be used. A per-candidate acceptance function can be extracted from each data event itself and included in the likelihood by appropriately adjusting the intervals over which normalisation is performed; this procedure was pioneered by the ACCMOR collaboration [82], and refined by the CDF collaboration [83, 84]; it is called swimming. From the primary and secondary vertex topology of a

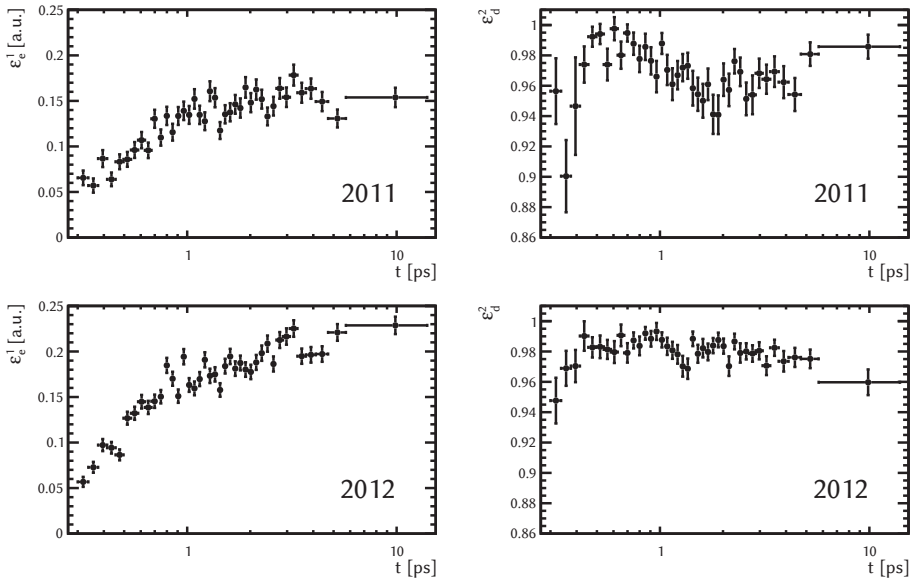


Figure 5.2: Result of the final likelihood fit for ϕ_s, including a binned acceptance with floating values, for the HLT1 single track, ε_e, and HLT2 detached dimuon, ε_d, selections for the 2011 and 2012 samples.

given event the decay-time interval in which this event would pass the HLT selection requirements can be determined. The acceptance function takes the form of one or more decay-time intervals in which a candidate would be accepted.

5.3.1 Background and Justification

The trigger of the CDF detector contained a dedicated component implemented in hardware, called the Silicon Vertex Tracker (SVT), which found tracks and vertices and applied selection requirements on the minimal and maximal impact parameters of found tracks [85, p. 13]. The SVT significantly improved the efficiency of identifying events which contained B-meson decays at trigger level in CDF, but the selection requirements on the impact parameters resulted in a complicated decay-time dependent acceptance. Since large differences were observed between data and available simulated samples, the data-driven swimming method was used to include the decay-time dependent acceptance in the likelihood fit. The following shows — in the case of perfect detector resolution — how this is achieved.

Starting from the PDF to observe a decay at time $0 < t_i < \inf$ with perfect resolution in an exponential distribution, the log-likelihood is given by:

$$\mathrm{PDF}(t_i) = \frac{1}{\tau} \mathrm{e}^{-t_i/\tau}, \tag{5.9}$$

$$\mathcal{L} = \prod_{i=1}^{N} \mathrm{P}(t_i), \tag{5.10}$$

$$\log \mathcal{L} = -N \log \tau - \sum_{i=1}^{N} \frac{t_i}{\tau}. \tag{5.11}$$

If a selection requirements on minimal and maximal decay-time t_{min} and t_{max} are introduced, the PDF and likelihood become:

$$\begin{aligned} \mathrm{PDF}(t_i) &= \mathrm{PDF}(t_i | t_i \in [t_{\mathrm{min}}, t_{\mathrm{max}}]), \\ &= \frac{\frac{1}{\tau} \mathrm{e}^{-t_i/\tau}}{\int_{t_{\mathrm{min}}}^{t_{\mathrm{max}}} \frac{1}{\tau} \mathrm{e}^{-t'/\tau} dt'}, \\ &= \frac{\frac{1}{\tau} \mathrm{e}^{-t_i/\tau}}{\mathrm{e}^{-t_{\mathrm{min}}/\tau} - \mathrm{e}^{-t_{\mathrm{max}}/\tau}}, \end{aligned} \tag{5.12}$$

$$\log \mathcal{L} = -N \log \tau - N \log \left(\mathrm{e}^{-t_{\mathrm{min}}/\tau} - \mathrm{e}^{-t_{\mathrm{max}}/\tau} \right) - \sum_{i=1}^{N} \frac{t_i}{\tau}. \tag{5.13}$$

If the decay-time selection requirements t_{min} and t_{max} differ per candidate, but are independent of the decay time t_i, the central term in eq. (5.13) becomes a sum over all candidates.

$$\log \mathcal{L} = -N \log \tau - \sum_{i=1}^{N} \left(\log \left(\mathrm{e}^{-t_{\text{min},i}/\tau} - \mathrm{e}^{-t_{\text{max},i}\tau} \right) \right) - \sum_{i=1}^{N} \frac{t_i}{\tau}. \tag{5.14}$$

For each candidate, the acceptance intervals are determined by re-evaluating the selection requirements across the full range of possible candidate decay-times, without changing its kinematic properties, and determining at which decay-time a candidate would still be accepted. This requires modification of a candidate's decay-time in the event, which would ideally be accomplished by moving a candidate's detector hits such that the re-reconstructed candidate would have a larger or smaller decay-time. The selections would then be re-evaluated on the modified event to determine if a candidate would still be selected. Due to the discrete nature of the tracking detectors, modifying the position of hits is very difficult and attempting it would likely introduce more uncertainty than the whole procedure is designed to remove.

Since the decay-time is proportional to the distance between a candidate's assigned primary vertex and its decay vertex, a candidate's decay-time can also be modified by translating the primary vertex' position along the direction of the momentum of the candidate. Once the primary vertex' position has been changed, selection requirements are re-evaluated to determine if a candidate would still be selected. This is illustrated in fig. 5.3 for the case of a selection requirement of a single daughter track's impact parameter in the presence of two primary vertices along the direction of the momentum of a candidate.

5.3.2 Swimming Algorithm

The LHCb L0-trigger does not contain any requirements on observables which are correlated with the decay-time of candidates. The complete implementation in software of the HLT facilitates calculation of the acceptance intervals, because the software trigger selections are easily re-evaluated.

The purpose of the algorithm which extracts the acceptance intervals is to identify those decay-times at which a selection decision changes. These so-called turning points are identified by translating all primary vertices along the direction of the momentum of a given candidate in steps from -200 to 200 mm, nearly half the length of the Velo. The steps are smaller when the translation produces candidate decay vertices close to $z = 0$, since that is the region where most

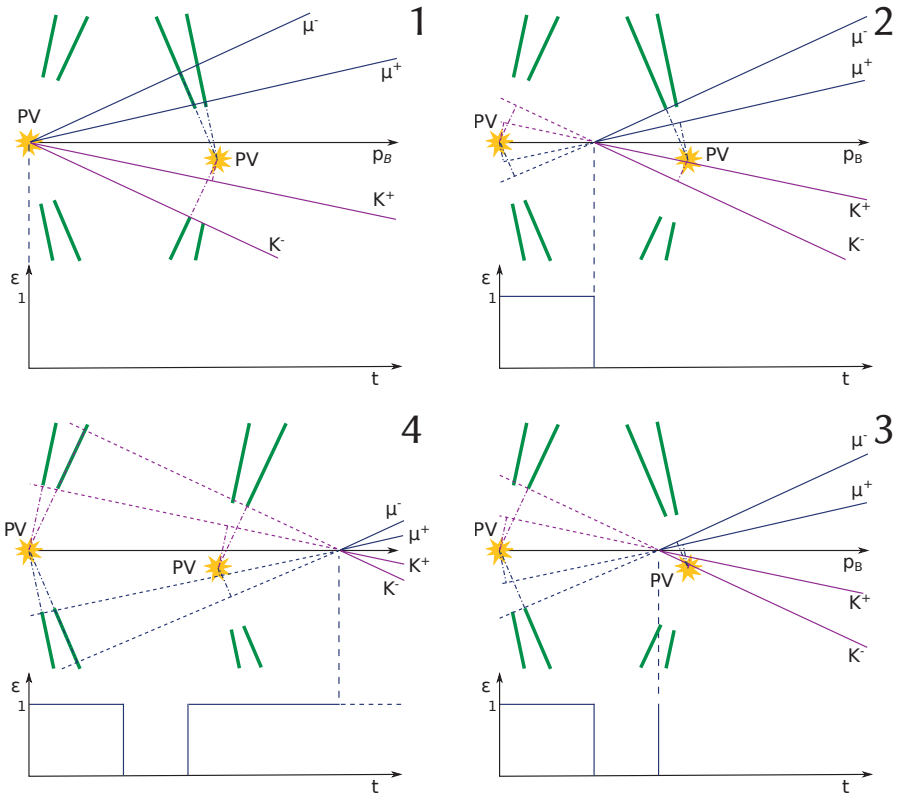


Figure 5.3: Illustration of the translation of a $B_s \to J/\psi\,\varphi$ candidate from it's primary vertex of origin to the end of the decay-time range. The green bars correspond to a selection requirement that the minimum impact parameter — indicated by the dashed-dotted lines — with respect to any primary vertex, of at least one track, is larger than a given value. If the requirement if satisfied — indicated by at least one impact parameter overlapping with it's corresponding green bar — the acceptance is 1 and 0 otherwise.

primary vertices are located. The number of and the distance between steps has been optimised to minimise the number of steps required while still being able to identify all turning points.

Once a decision is observed to change between two steps, a turning point is assumed to be present and the location of the two adjacent steps is saved. All turning points identified in the initial sweep are refined by re-evaluating the selections at smaller steps between the points where the decision is observed to change. The decay-time of a candidate is recalculated at each step and saved alongside the translation distance for each turning point. Since the decay-time is calculated in the same way as the decay-time used in the likelihood fit, no further corrections are required.

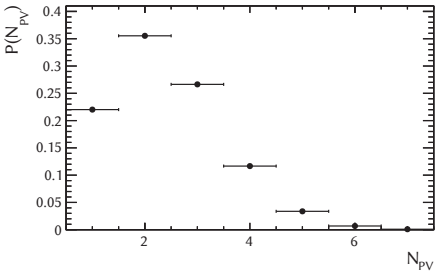
The procedure is repeated for all trigger selections of interest. The turning points are combined according to which trigger selections are required by the final selection. If the final selection itself contains requirements on decay-time correlated quantities, the swimming procedure can be applied to it too and the final set of turning points is obtained from the logical AND of the turning points for the trigger and the final selections.

5.3.3 Swimming Validation

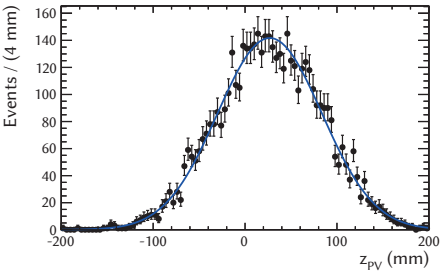
To validate the swimming procedure a number of fast simulations are performed, in which acceptance intervals are generated from a simplified arrangement of a number of primary vertices along the z-axis. The generated acceptance intervals are used to generated candidate decay-times from a single exponential distribution. The distribution is fit with an exponential distribution, which is normalised per event according to the generated acceptance intervals as defined by eq. (5.13).

To generate the acceptance intervals the probability for the number of primary vertices in an event $\mathrm{P}(N_{\mathrm{PV}})$ and the distribution of primary vertices along the z-axis are taken from data. $\mathrm{P}(N_{\mathrm{PV}})$ is shown in fig. 5.4a and the distribution of primary vertices along the z-axis is obtained from a minimum bias sample and shown in fig. 5.4b. Primary vertices are observed to be distributed along the z-axis according to a Gaussian distribution with a mean of 27 mm and a with of 57 mm from a fit to the position of primary vertices in L0 accepted data.

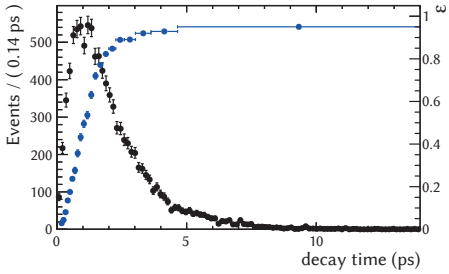
Using the obtained distributions, the number of primary vertices and their position along the z-axis is generated. Around each generated vertex, an exclusion zone is generated according to a Gaussian distribution with an empirically chosen mean of 2 mm and a width of 3 mm; a single number d is generated from this Gaussian and the exclusion zone is taken to be $(z_{\mathrm{PV}} - |d|, z_{\mathrm{PV}} + |d|)$ mm, for each vertex. One of the primary vertices is randomly selected to be the ori-



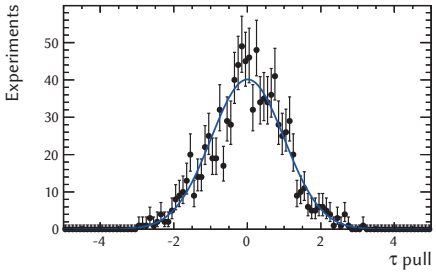
(a) Probability for an event to contain a number of primary vertices for the 2011 calibration sample.



(b) Distribution of primary vertices along the z-axis in run 130560 (2012).



(c) Generated decay-time distribution of one of the fast simulations. The event generation efficiency is superimposed in blue.



(d) Pull distribution of the fitted lifetime τ after 1000 fast simulations.

Figure 5.4: Several distributions related to the validation of the per-event decay-time acceptance.

gin vertex of the B_s meson and using this vertex, the decay-times corresponding to all exclusion zones are calculated according to eq. (4.2), assuming a typical B_s candidate momentum of 79 GeV/c. The acceptance intervals are finally obtained by subtracting the exclusion zones from the decay-time range of $[0, 14]$ ps, where 14 ps is the maximum decay-time used. Once the acceptance intervals are known, the decay-time distribution is generated such that for each event the decay-time is inside one of the generated acceptance intervals. The resulting decay-time distribution is shown in fig. 5.4c, where the effect of the acceptance is clearly visible at low decay-times.

A likelihood fit of a single exponential is performed to the generated decay-time distribution and the pull distribution for the lifetime τ of the exponential distribution after 1000 fast simulations is shown in fig. 5.4d. A fit of a Gaussian to the pull distribution resulted in a mean of 0.008 ± 0.031 and a width of 0.994 ± 0.022, which indicate that no biases are incurred by using decay-time acceptance intervals in the fit and that the uncertainty on the fitted lifetime using the acceptance method is correctly estimated.

Chapter 6

Decay-Time Resolution

The presence of a finite decay-time resolution smears out the observed decay-time value with respect to its true value, which in turn results in a dampening of the observable amplitude of flavour oscillations and time-dependent CP-asymmetries. This effect is taken into account in the PDF by convoluting all decay-time dependent terms of the physics PDF with the model for the decay-time resolution. To avoid wrongly attributing a smaller observed oscillation amplitude to a physics effect, the decay-time resolution model should be carefully calibrated.

The uncertainties on the position and direction of reconstructed tracks result in uncertainties on the position of primary and secondary vertices, which are required for reconstruction of a candidate's decay-time as can be seen from eq. (4.1). If the uncertainty on the measured decay-time is known per candidate, including it in the likelihood fit improves its performance. As described in section 4.1.2, the kinematic fit, which is used to extract the decay-time of a candidate, also provides an estimate of its uncertainty, the estimated decay-time error, σ_t. Since σ_t is not guaranteed to be identical to the true decay-time error, $\sigma_{t,\text{true}}$, it must be calibrated. To that end it is assumed that $\sigma_{t,\text{true}}$ can be obtained from a polynomial in σ_t. The choice of parameterisation, the extraction of its parameters from data and the estimation of the corresponding systematic uncertainties are the main subjects of this chapter.

6.1 Calibration Procedure

The decay-time resolution causes about half of the reconstructed decay-times to be shorter than their true values and, for short decay-times, the reconstructed

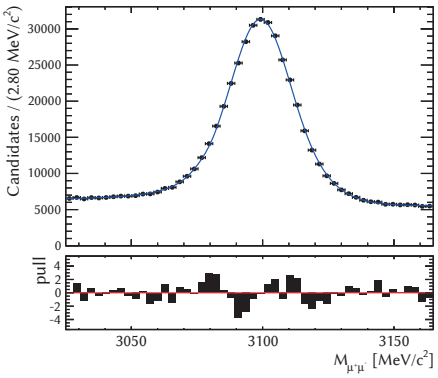


Figure 6.1: Distribution of $M_{\mu^+\mu^-}$ for the 2011 calibration sample.

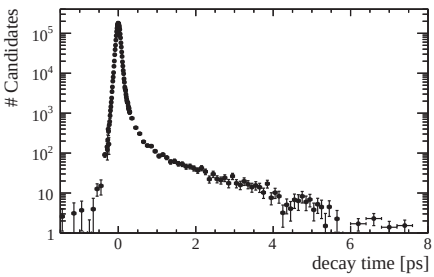


Figure 6.2: Decay-time distribution of prompt-background candidates in 2011 data, weighted using sWeights that have been obtained from a fit to the distribution of $M_{\mu^+\mu^-}$.

decay-time may therefore be negative. Since the decay-time resolution is the only source of negative decay-times, the negative time component of the decay-time distribution may be used to calibrate the resolution model.

However, the selection criteria of the so-called detached sample prevent such a procedure. Alternatively, candidates are required to have been selected by the so-called prescaled selection which, by construction, has a decay-time independent efficiency. The prescale factor applied in the prescaled selection results in only approximately three thousand signal candidates being accepted, which is insufficient to extract the resolution model parameters with sufficient precision using signal candidates.

To obtain a precise calibration, background candidates that include a $J/\psi \to \mu^+\mu^-$ decay are used instead. Most J/ψ mesons are not created as a result of B-decays, but directly in the strong interaction processes of a proton-proton interaction. The lifetime of the J/ψ is so short that a B_s candidate created from such a prompt J/ψ and two kaons which also originated from strong interaction process has a true decay-time consistent with 0 ps. Such B_s candidates are called prompt candidates. Calibration of the resolution model with prompt candidates is described in section 6.2. The calibration proceeds in three steps. A fit with a Gaussian resolution model, with a width given by a constant scale factor times the estimated decay-time uncertainty, σ_t, is performed in section 6.2.1 to provide an example and frame of reference for the subsequent improvements. Since a fit with a double Gaussian distribution results in large correlations between the

parameters describing the widths of the Gaussians, a reparameterisation is implemented in section 6.2.2 to reduce the unwanted correlations. The final parameterisation of σ_t is determined in section 6.2.3, where the values of the parameters for the final fit are also listed.

To estimate the systematic uncertainties related to the use of prompt-background candidates instead of the signal candidates to calibrate the resolution model, the calibration procedure is repeated in section 6.3 with simulated samples of prompt-background and signal candidates and the calibrations are compared to obtain the systematic uncertainties listed in section 6.3.4. A parameterisation of the mean of the Gaussian model is determined to be required for simulated signal candidates in section 6.3.2 and the effect of momentum is studied in section 6.3.3.

A the end of the chapter, primary vertex refitting and primary vertex assignment are studied in section 6.4 and section 6.5, respectively.

6.2 Calibration on Data

The sample of prompt-background candidates contains B_s candidates created from two random muons and two random kaons in addition to those created from a J/ψ and two random kaons. Since the kinematic properties of combinatorial background candidates are less similar to signal $B_s \to J/\psi\varphi$ candidates, combinatorial background candidates are not used to calibrate the decay-time resolution model. The decay-time distribution of the combinatorial background component is statistically subtracted from the full sample by fitting the $\mu^+\mu^-$ invariant-mass distribution and using the result of the fit to calculate sWeights, as described in section 4.2. The PDF used to fit this sample is given by:

$$\mathrm{PDF}(M) = N_{J/\psi} \times ((f_{\mathrm{CB}}\mathrm{CB}_1(M) + (1 - f_{\mathrm{CB}})\mathrm{CB}_2(M)) + N_{\mathrm{bkg}} \times \mathrm{e}^{-M/c}, \quad (6.1)$$

where $M = M_{\mu^+\mu^-}$, CB_1 and CB_2 are Crystal-Ball PDFs, and the PDF includes an overall normalisation, not explicitly mentioned here. The Crystal Ball PDF [58, 59] is a Gaussian PDF with mean μ and width σ with a power-law tail ($n = 2$, in this case) on one side (the left side, in this case). The boundary between the power-law and Gaussian PDFs is defined by $(m - \mu)/\sigma > -\alpha$. The result of a fit to the 2011 data is shown in fig. 6.1. The sWeights are then used in a consecutive weighted-maximum-likelihood fit to extract the parameters of the resolution model. Figure 6.2 shows the weighted distribution of decay-times; the peak around 0 ps originates from prompt-background candidates, while the tail at higher decay-time consist of J/ψ candidates from various $B^0 \to J/\psi X$ decays.

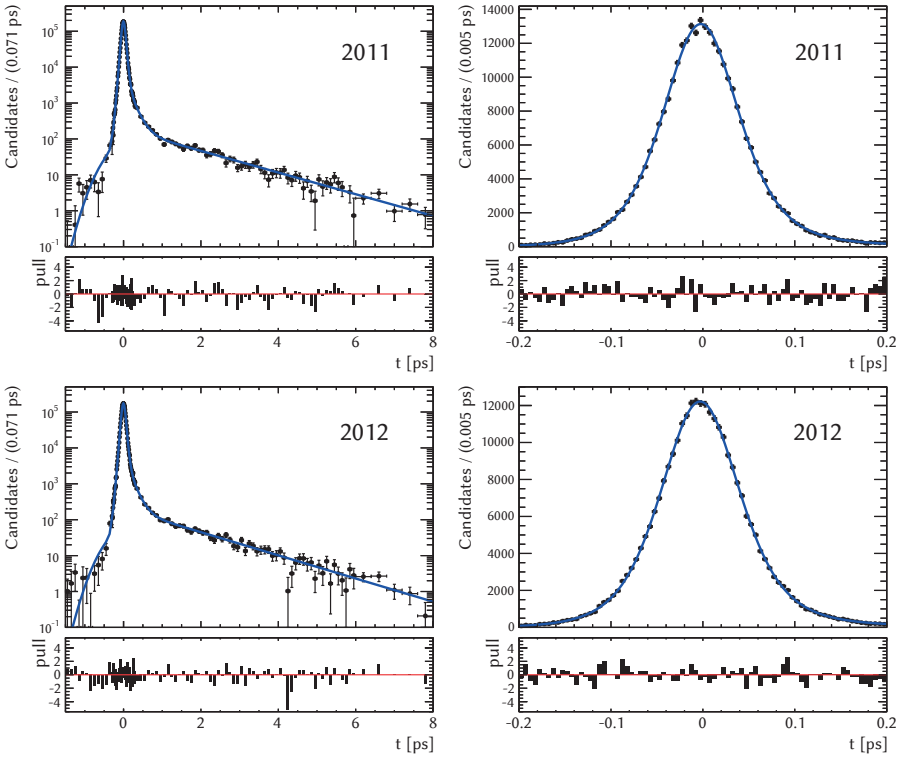


Figure 6.3: Left: results of fits to the decay-time distribution of the 2011 and 2012 sample of prompt-background candidates shown on a log-scale. Right: a linear scale zoomed-in around 0. Below the distributions the pull of the PDF with respect to the data points is shown, demonstrating the good quality of the fit.

6.2.1 Fit with Constant Scale-Factors

To give an idea of how the resolution model is calibrated on a sample of prompt background events, the PDF that is fitted to this data sample is described, followed by the results of a fit to the 2011 sample.

The component of the prompt-background sample, that peaks around the J/ψ invariant mass, does not have a 100% purity of prompt candidates: long-lived candidates and candidates associated to the wrong PV are also present. The PDF fitted to the prompt-background sample thus consists of three components, two of which are convoluted with the resolution model: a Dirac delta function for

parameter	value	
N_{prompt}	289 726.0	$\pm$ 63.9
N_{ll}	15 093.0	$\pm$ 35.7
N_{WPV}	850.0	$\pm$ 7.5
f_{sl}	0.7457	$\pm$ 0.0071
τ_{ll}	1.428	$\pm$ 0.032 [ps]
τ_{sl}	0.1433	$\pm$ 0.0044 [ps]
f_2	0.241	$\pm$ 0.012
μ	-2.598	$\pm$ 0.098 [fs]
s_1	1.2071	$\pm$ 0.0076
s_2	2.089	$\pm$ 0.027
σ_{WPV}	0.377	$\pm$ 0.017 [ps]

Table 6.1: Result of a fit to the 2011 sample of prompt-background candidates.

the prompt peak and the two exponentials to cover the long-lived background. The third component is a Gaussian PDF, with mean 0 and width σ_{WPV}, which models the contribution from candidates that were assigned a wrong PV. Since the resolution model is not expected to provide information for such candidates, the Gaussian PDF that models them is not convoluted with it. The PDF is given by:

$$\begin{aligned}\text{PDF}(t|\sigma_t) \propto &\left[(1-f_2)\mathrm{e}^{-\frac{1}{2}\left(\frac{t-\mu}{s_1\sigma_t}\right)^2} + f_2\mathrm{e}^{-\frac{1}{2}\left(\frac{t-\mu}{2s_2\sigma_t}\right)^2}\right] \\ &\otimes \left[N_{\text{prompt}}\delta(t) + N_{\text{ll}}\left(f_{\text{sl}}\mathrm{e}^{-t/\tau_{\text{sl}}} + (1-f_{\text{sl}})\mathrm{e}^{-t/\tau_{\text{ll}}}\right)\right] \\ &+ \; N_{\text{WPV}}\mathrm{e}^{-\frac{1}{2}\left(\frac{t}{\sigma_{\text{WPV}}}\right)^2}\end{aligned} \tag{6.2}$$

where A_{sl} and A_{ll} are normalisation constants. Since the resolution model uses the per-candidate reconstructed decay-time error, σ_t, it is assumed that the true per-candidate decay-time error, $\sigma_{t,\text{true}}$, can be obtained from a parameterisation utilising the estimated error. In this section, two scale factors have been chosen to obtain $\sigma_{t,\text{true}}$ from σ_t, represented by s_1 and s_2.

A fit of this PDF to the 2011 sample of prompt-background events was performed and the results are listed in table 6.1.

From the residual distributions shown in fig. 6.3 it can be seen that the PDF fits the data well. The value obtained for the decay-time of the long-lived background component τ_{ll} listed in table 6.1 is very close to the decay-time of B^0 and

	f_2	s_1	s_2
f_2	1	−0.935	−0.935
s_1	−0.935	1	0.834
s_2	−0.935	0.834	1

(a) Before reparameterisation.

	f_2	$\overline{\sigma}$	σ_σ
f_2	1	−0.237	−0.304
$\overline{\sigma}$	−0.237	1	0.476
σ_σ	−0.304	0.476	1

(b) After reparameterisation.

Table 6.2: Correlations between the parameters describing the widths of the two Gaussian components of the resolution model before and after reparameterisation according to eq. (6.3).

B_s mesons, which indicates that this component consists of their decays. The fact that the common mean of the Gaussians, μ, is significantly different from zero indicates that there is a systematic bias in the reconstructed decay-times.

6.2.2 Reparameterisation of the Resolution Model

When the maximum-likelihood fit of the model of eq. (4.13) to the 2011 prompt-background sample is performed to obtain the parameters of the resolution model, large correlations appear between the widths of the Gaussian components and the fraction of each Gaussian. The correlation is listed in table 6.2a.

To improve the convergence of the fit and reduce these correlations, a new parameterisation has been introduced, which is effectively a weighted sum and difference of the parameters and is described in more detail in appendix D:

$$\begin{aligned} \overline{\sigma} &= (1-f)\sigma_1 + f\sigma_2 \\ \sigma_\sigma &= \sqrt{f(1-f)}\,(\sigma_1 - \sigma_2)\,, \end{aligned} \tag{6.3}$$

The inverse transform is:

$$\begin{aligned} \sigma_1 &= -\sqrt{\frac{f}{1-f}}\sigma_\sigma + \overline{\sigma} \\ \sigma_2 &= \sqrt{\frac{1-f}{f}}\sigma_\sigma + \overline{\sigma}, \end{aligned} \tag{6.4}$$

Since this reparameterisation is applied to the widths of the two Gaussian components of the decay-time resolution model, it is independent of any parameterisation of σ_t. A fit of this model to the 2011 sample yields the results shown in

parameter	value 2011	value 2012
N_{prompt}	289 737.0 $\pm$ 63.9	290 610.0 $\pm$ 66.5
N_{ll}	15 092.0 $\pm$ 35.7	16 538.0 $\pm$ 43.0
N_{WPV}	844.0 $\pm$ 6.3	675.0 $\pm$ 6.2
f_{sl}	0.7458 $\pm$ 0.0070	0.7785 $\pm$ 0.0065
τ_{ll}	1.429 $\pm$ 0.032 [ps]	1.295 $\pm$ 0.031 [ps]
τ_{sl}	0.1433 $\pm$ 0.0044 [ps]	0.1358 $\pm$ 0.0043 [ps]
f_2	0.241 $\pm$ 0.012	0.228 $\pm$ 0.011
μ	$-$2.598 $\pm$ 0.098 [fs]	$-$3.34 $\pm$ 0.11 [fs]
$\overline{\sigma}$	1.4198 $\pm$ 0.0027	1.4792 $\pm$ 0.0030
σ_σ	0.3774 $\pm$ 0.0047	0.4077 $\pm$ 0.0051

Table 6.3: Fit results for the 2011 and 2012 prompt-background samples.

table 6.2b. A clear reduction in the correlation of the parameters of the resolution model is observed.

The results of maximum-likelihood fits of the PDF given in eq. (6.2) to the 2011 and 2012 samples of prompt-background candidates are listed in table 6.3 and shown in fig. 6.3.

6.2.3 Calibration of the Per-Event Error Parameterisation

The parameterisation of the widths of the two Gaussian components of the decay-time resolution model is expected to be a function of the reconstructed resolution, σ_t, itself and it is calibrated by performing a simultaneous fit to the prompt sample in bins of σ_t. The yields of the different components of the PDF — prompt, long-lived background and wrong-PV background — are separately determined for each bin, as are the parameters $\overline{\sigma}$ and σ_σ. The fraction of the second Gaussian is observed to be constant across the bins within errors, so a single parameter is shared among the bins; the common mean, μ of the Gaussians is also shared. Since this PDF is binned in σ_t, the PDFs for the individual bins no longer depend on σ_t. The result of such fits to the 2011 and 2012 samples are shown as the data points in fig. 6.4; the histogram in the background is constructed such that the area of a bin represents the fraction of the candidates which falls in that bin.

Performing the final fit in bins of σ_t is not desirable due to the large additional level of complexity this would entail. Determining the systematic errors resulting from the calibration procedure would also become more complicated.

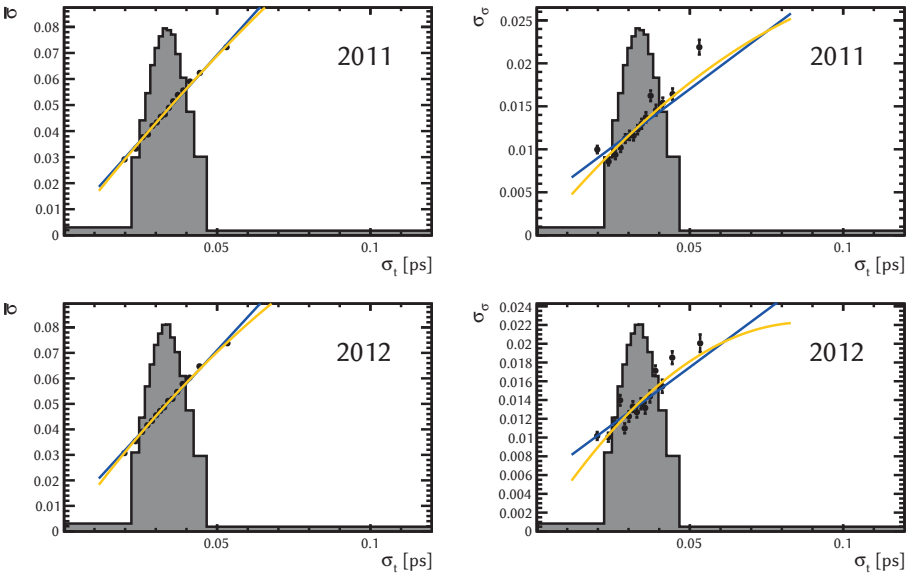


Figure 6.4: Result of a simultaneous fit in bins of σ_t to the 2011 (top) and 2012 (bottom) sample of prompt-background candidates. The grey histogram represents the fraction of candidates in each bin. The result of the simultaneous fit in bins of σ_t is shown as the data points, while the results of fits with a parameterisation of σ_t directly included are shown as coloured lines: linear (blue), quadratic without offset (orange).

year	$\overline{\sigma_t}$
2011	0.0350
2012	0.0349

Table 6.4: Average estimated decay-time resolutions.

To simultaneously solve these problems, a parameterisation of $\overline{\sigma}$ and σ_σ as a function of σ_t has directly been introduced into the calibration fits. Since the decay-time resolution model consists of two independent Gaussian components, the parameterisations of $\overline{\text{sf}}$ and sf_σ as a function of σ_t do not share parameters. An additional advantage of parameterising $\overline{\text{sf}}$ and sf_σ as a function of σ_t is that these dependencies can be studied in simulated samples and any differences can be translated into a systematic error.

The reparameterisation of the widths of the two Gaussians as introduced in eq. (6.4) is independent of the parameterisations of $\overline{\sigma}$ and σ_σ as a function of σ_t, as long as these are applied before eq. (6.4). To limit the number of parameters which are added to the fit, two possible parameterisations have been selected, called "linear" and "quadratic":

$$\sigma_{t,\text{true}} = c_0' + c_1'(\sigma_t - \overline{\sigma_t}), \tag{6.5}$$

$$\sigma_{t,\text{true}} = c_0(\sigma_t - \overline{\sigma_t}) + c_1(\sigma_t - \overline{\sigma_t}))\sigma_t, \tag{6.6}$$

where $\overline{\sigma_t}$ is the mean estimated decay-time error for the sample under consideration; table 6.4 lists the appropriate values. Figure 6.4 shows the result of the two different parameterisations having been inserted in the calibration fit as lines with blue, and orange colour. It can be seen that both the linear parameterisation and the quadratic parameterisation without offset describe the widths of the Gaussian components well.

The true decay-time resolution is expected to be zero when the estimated decay-time resolution is zero and since the quadratic parameterisation without offset has this feature, it is preferred over the linear parameterisation. The result of a fit to the calibration samples with this parameterisation included is listed in table 6.5.

6.3 Cross-checks with Simulated Candidates

Since the calibration of the decay-time resolution-model is performed using prompt-background candidates as a proxy for signal candidates, a systematic uncertainty

parameter	value 2011	value 2012
N_{prompt}	290 012.0 $\pm$ 63.3	291 308.0 $\pm$ 66.2
N_{ll}	14 798.0 $\pm$ 34.6	15 770.0 $\pm$ 39.1
N_{WPV}	857.0 $\pm$ 7.5	745.0 $\pm$ 8.4
f_{sl}	0.7436 $\pm$ 0.0072	0.7736 $\pm$ 0.0068
τ_{sl}	0.1479 $\pm$ 0.0045 [ps]	0.1447 $\pm$ 0.0044 [ps]
τ_{ll}	1.438 $\pm$ 0.032 [ps]	1.315 $\pm$ 0.031 [ps]
$\overline{\sigma}$: c_0	1.4207 $\pm$ 0.0027	1.4811 $\pm$ 0.0030
$\overline{\sigma}$: c_1	$-$2.85 $\pm$ 0.27	$-$4.96 $\pm$ 0.28
f_2	0.244 $\pm$ 0.012	0.239 $\pm$ 0.012
σ_σ: c_0	0.3778 $\pm$ 0.0048	0.4057 $\pm$ 0.0053
σ_σ: c_1	$-$1.55 $\pm$ 0.49	$-$2.88 $\pm$ 0.52
μ	$-$2.55 $\pm$ 0.10 [fs]	$-$3.22 $\pm$ 0.10 [fs]
σ_{WPV}	0.377 $\pm$ 0.017 [ps]	0.358 $\pm$ 0.019 [ps]

Table 6.5: Result of a fit to the 2011 and 2012 samples of prompt-background candidates, where a quadratic parameterisation without offset is introduced into the fit.

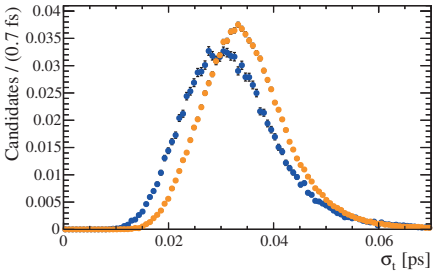


Figure 6.5: σ_t distributions for the 2012 calibration (orange) and signal (blue) samples.

year	$\overline{\sigma_t}_{,\text{sig}}$	$\overline{\sigma_t}_{,\text{prompt}}$
2011	0.0332	0.0360
2012	0.0331	0.0367

Table 6.6: Average estimated decay-time resolutions for simulated samples.

should be assigned to the assumption that a calibration with prompt-background candidates yields the same result as a calibration with signal candidates. Figure 6.5 shows that significant differences exist between the distributions of the reconstructed per-candidate decay-time error, σ_t, of prompt-background and signal candidates. Since the σ_t of the signal candidates is used in the per-event likelihood fit and the calibration of the resolution-model does not depend on the distribution of σ_t, no a priori systematic error needs to be assigned due to the different distributions.

The systematic uncertainty corresponding to the calibration of the parameters of the resolution-model with prompt-background candidates is evaluated using simulated events. To do this, the selection described in section 4.1.4 is applied to a sample of approximately 40 million simulated events containing a J/ψ decaying to two muons to obtain a sample of simulated prompt-background candidates.

The same fit as used for the prompt-background sample in data (eq. (6.2)) is run on this simulated sample to obtain a calibration of the resolution-model. The only difference is that the value of $\overline{\sigma_t}$ for the sample fit is used. The values for samples of prompt-background and signal candidates simulated with 2011 and 2012 conditions are given in table 6.6. Subsequently, a sample of simulated $B_s \to J/\psi\varphi$ signal candidates is directly used to obtain an alternative calibration of the resolution-model. The two different sets of parameters are then used in the final fit and the observed differences in the physics parameters of interest are taken as systematic uncertainties of the calibration. The fits performed to obtain these two sets of calibration parameters are described in the following section.

All prompt-background candidates used for calibration have $t = 0$ by construction and from eq. (4.12) it can be seen that the candidate's momentum uncertainty does not contribute to σ_t for $t = 0$. The calibration sample used for the calibration of σ_t can therefore not be used to calibrate the contribution of the momentum uncertainty to σ_t. A test of the assumption that the estimated decay-time error takes the momentum uncertainties into account properly, or in other words, that no extra calibration as a function of momentum is required, is described in section 6.3.3.

parameter	value 2011			value 2012		
N_{prompt}	31 168.0	$\pm$ 21.3		30 321.0	$\pm$ 22.1	
N_{ll}	3011.0	$\pm$ 12.7		2813.0	$\pm$ 13.5	
N_{WPV}	148.0	$\pm$ 3.3		180.0	$\pm$ 4.9	
f_{sl}	0.609	$\pm$ 0.019		0.620	$\pm$ 0.020	
τ_{sl}	0.144	$\pm$ 0.011		0.139	$\pm$ 0.011	
τ_{ll}	1.450	$\pm$ 0.057		1.489	$\pm$ 0.060	
$\overline{\sigma}$: c_0	1.2463	$\pm$ 0.0076		1.315	$\pm$ 0.011	
$\overline{\sigma}$: c_1	0.03	$\pm$ 0.65		-1.40	$\pm$ 0.71	
f_2	0.235	$\pm$ 0.053		0.135	$\pm$ 0.040	
σ_σ: c_0	0.297	$\pm$ 0.015		0.346	$\pm$ 0.026	
σ_σ: c_1	-4.1	$\pm$ 1.4		-2.9	$\pm$ 1.7	
μ	-0.72	$\pm$ 0.27	[fs]	-0.56	$\pm$ 0.30	[fs]
σ_{WPV}	0.322	$\pm$ 0.037		0.381	$\pm$ 0.047	

Table 6.7: Results of a fit to a sample of simulated events containing a $J/\psi \to \mu^+\mu^-$ decay. $B_s \to J/\psi\varphi$ candidates have been created by by applying the prescaled selection. The parameters c_0 and c_1 are defined in eq. (6.6), while $\overline{\sigma}$ and σ_σ are defined in eq. (6.3).

6.3.1 Fits to Simulated Samples

For prompt-background candidates, the true decay-time, t_{true}, is 0, so the distributions of t and $t - t_{\text{true}}$ are identical. Since the simulated prompt-background sample contains also some B-decays of various types, some longer lived background candidates are expected, as in data. It is for this reason that the same PDF as used for data (eq. (6.2)) is fitted to this sample. The results of this fit to the samples simulated with 2011 and 2012 conditions are listed and shown in table 6.7 and fig. 6.6. Comparing table 6.7 with the values obtained from the fits to data listed in table 6.5, several differences are observed: the values of $\overline{\sigma}$: c_0 are smaller for simulated prompt-background samples, which indicates that for the simulation σ_t is closer to $\sigma_{t,\text{true}}$ on average than it is in data;

For the simulated sample containing $B_s \to J/\psi\varphi$ signal decays, asymmetric tails are observed in the distribution of $t - t_{\text{true}}$ shown as the data points in fig. 6.8. To understand how to model these tails, the distributions of $(t - t_{\text{true}})/\sigma_t$ for samples simulated with 2011 and 2012 conditions are shown as the data points in fig. 6.7. On the log-scale of these figure, it can be seen that the tails are approximately straight lines, indicating they can be modelled by exponential PDFs. To

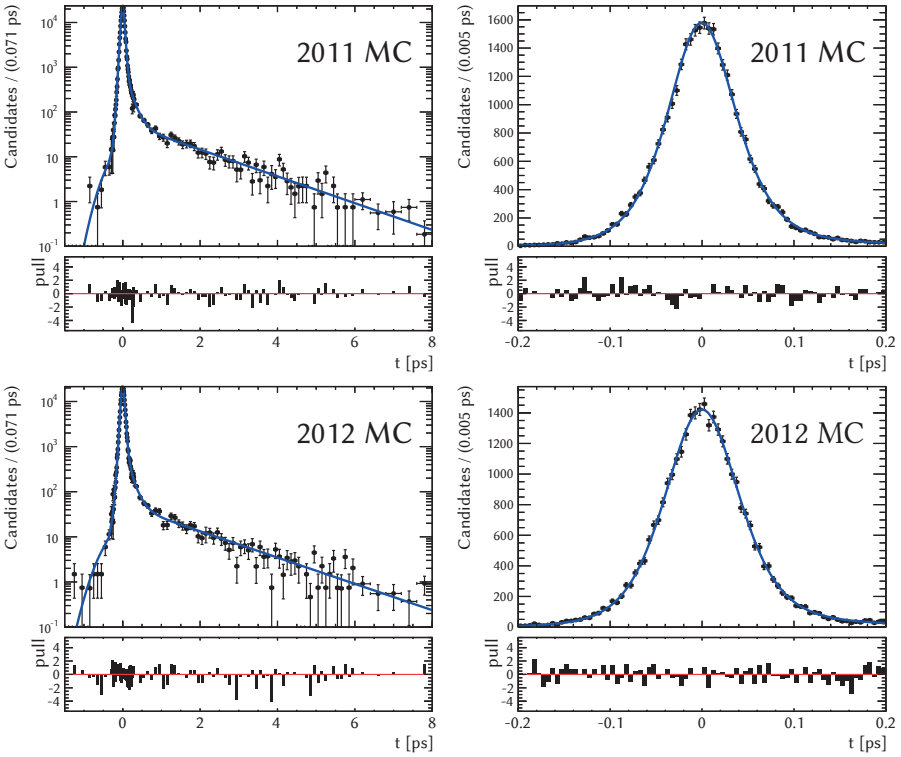


Figure 6.6: Left: result of a likelihood fit to the decay-time distribution of selected $B_s \to J/\psi\varphi$ candidates obtained from the 2011 and 2012 simulated prompt-background samples. The fitted curve is the PDF described in eq. (6.2). Right: zoomed-in view around $t = 0$.

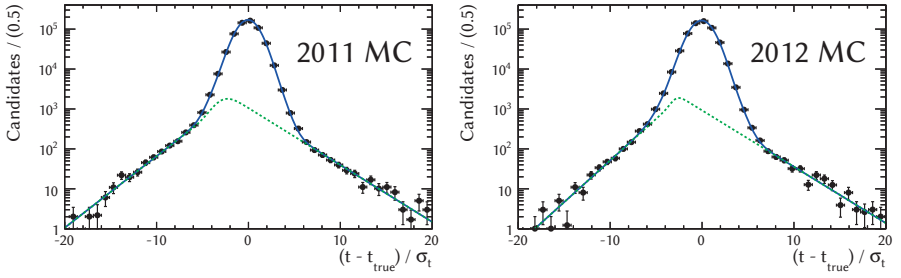


Figure 6.7: Results of a likelihood fit of the PDF defined in eq. (6.9) to the distribution of $(t - t_{\text{true}})/\sigma_t$ of selected $B_s \to J/\psi\varphi$ candidates obtained from samples simulated with 2011 and 2012 conditions

allow the tails to be of different size and to achieve a smooth transition between them, two PDFs are used, which are both Gaussians convoluted with an exponential on one side. This PDF is given by:

$$\mathrm{erfc}(t') = \frac{2}{\sqrt{\pi}} \int_{t'}^{\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x, \tag{6.7}$$

$$a = \frac{\sigma^2}{2\tau^2} \pm \frac{t'}{\tau},$$

$$\mathrm{GE}(t') = 2\tau \mathrm{e}^{a} \times \mathrm{erfc}\left(\frac{\sigma}{\sqrt{2}\tau} \pm \frac{t'}{\sqrt{2}\sigma}\right), \tag{6.8}$$

where t' is $(t - t_{\mathrm{true}})/\sigma_t$, $\mathrm{erfc}(t')$ is the complimentary error function, σ is the width of the Gaussian side and τ the lifetime of the exponential side. For one of the components, the exponential part is on the left side, while for the other it is on the right side.

The PDF fit to the distributions of $(t - t_{\mathrm{true}})/\sigma_t$ shown in fig. 6.7 becomes:

$$\begin{aligned}\mathrm{PDF}(t - t_{\mathrm{true}}) = {} & (1 - f_{\mathrm{T}})\left[(1 - f_{\mathrm{G}})\mathrm{G}_1(t') + f_{\mathrm{G}}\mathrm{G}_2(t')\right] \\ & + f_{\mathrm{T}}\left[(1 - f_{\mathrm{E}})\mathrm{GE}_1(t') + f_{\mathrm{E}}\mathrm{GE}_2(t')\right],\end{aligned} \tag{6.9}$$

where G_1 and G_2 are Gaussian PDFs and GE_1 and GE_2 are defined by eq. (6.8).

To adapt this model to allow it to be fit to the distribution of $t - t_{\mathrm{true}}$, instead of $(t - t_{\mathrm{true}})/\sigma_t$, σ and τ of eq. (6.8) are replaced by $\sigma\sigma_t$ and $\tau\sigma_t$, respectively and the fractions by yields. The resulting PDF becomes:

$$\begin{aligned}\mathrm{PDF}(t - t_{\mathrm{true}}) = {} & N_{\mathrm{C}}\left[(1 - f_{\mathrm{G}})\mathrm{G}_1(t - t_{\mathrm{true}}|\sigma_t) + f_{\mathrm{G}}\mathrm{G}_2(t - t_{\mathrm{true}}|\sigma_t)\right] \\ & + N_{\mathrm{R}}\left[(1 - f_{\mathrm{E}})\mathrm{GE}_1(t - t_{\mathrm{true}}|\sigma_t) + f_{\mathrm{E}}\mathrm{GE}_2(t - t_{\mathrm{true}}|\sigma_t)\right],\end{aligned} \tag{6.10}$$

where the G_1 and G_2 are the Gaussian components of the resolution model as defined in eq. (4.13), GE_1 and GE_2 are defined by eq. (6.8), with $\sigma \to s_\sigma\sigma_t$ and $\tau \to s_\tau\sigma_t$ and s_σ and s_τ scale-factors.

After inclusion of the exponential tails, the PDF still somewhat overshoots the data, as can be seen on the right-hand side of the residual distributions in fig. 6.8. This may be caused by a feature of the PV refitting procedure applied during reconstruction of the signal candidates; this feature, and how to improve upon it, is discussed in section 6.4.

6.3.2 Parameterisation of the Common Mean, μ

In principle the reconstructed decay-time, t, might have a non-zero bias $t - t_{\mathrm{true}}$, which can depend on the reconstructed error σ_t. A simultaneous fit in

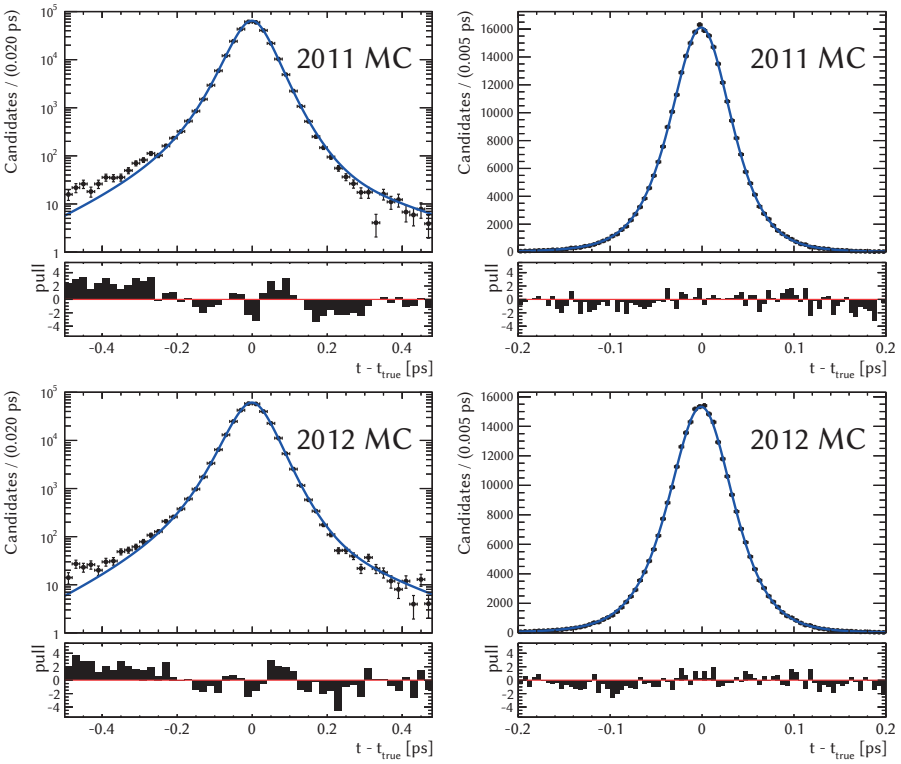


Figure 6.8: Left: result of a likelihood fit to the distribution of $t - t_{\text{true}}$ for samples simulated with 2011 and 2012 conditions containing a $B_s \to J/\psi\,\varphi$ decay. Right: zoomed-in view around $t - t_{\text{true}} = 0$.

parameter	value 2011	value 2012
N_{sig}	291 764.0 $\pm$ 68.6	290 138.0 $\pm$ 89.7
N_{tail}	3545.0 $\pm$ 42.7	5033.0 $\pm$ 72.1
$\overline{\sigma}$: c_0	1.2412 $\pm$ 0.0026	1.2795 $\pm$ 0.0035
$\overline{\sigma}$: c_1	1.96 $\pm$ 0.18	1.06 $\pm$ 0.18
f_2	0.180 $\pm$ 0.018	0.170 $\pm$ 0.024
σ: c_0	0.2512 $\pm$ 0.0057	0.2344 $\pm$ 0.0082
σ: c_1	3.14 $\pm$ 0.36	2.89 $\pm$ 0.39
μ_0	-1.49 $\pm$ 0.11 [fs]	-1.47 $\pm$ 0.13 [fs]
μ_1	-0.1706 $\pm$ 0.0082	-0.1819 $\pm$ 0.0087
μ_2	-6.28 $\pm$ 0.49 [ps^{-1}]	-6.91 $\pm$ 0.49 [ps^{-1}]

Table 6.8: Results of a fit to a sample of simulated events containing a $B_s \to J/\psi\varphi$ decay. $B_s \to J/\psi\varphi$ candidates have been created by applying the prescaled selection. The parameters describing the long tails have been omitted.

bins of σ_t to the samples of simulated events containing a $B_s \to J/\psi\varphi$ decay is performed, where for the common mean of the Gaussian components, μ, a separate parameter is used in each bin in σ_t. The fitted values for the means are shown as the data points in fig. 6.9, which indeed shows that μ depends on σ_t. To take this dependency into account, a linear and a quadratic parameterisation are considered and each is included in a fit of the PDF given in eq. (6.10). The parameterisations are given by:

$$\mu(\sigma_t) = \mu_0' + \mu_1'(\sigma_t - \overline{\sigma_t}), \tag{6.11}$$

$$\mu(\sigma_t) = \mu_0 + \mu_1(\sigma_t - \overline{\sigma_t}) + \mu_2(\sigma_t - \overline{\sigma_t})^2. \tag{6.12}$$

Simulated Signal Sample

For both 2011 and 2012 samples, the results of the fit including the quadratic parameterisation are listed in table 6.8. The χ^2/DoF of the quadratic parameterisation is for both samples significantly better than that of the linear parameterisation. The presence of the dependence of μ on σ_t means that the reconstructed decay-time, t is underestimated more for candidates with a larger decay-time uncertainty, σ_t.

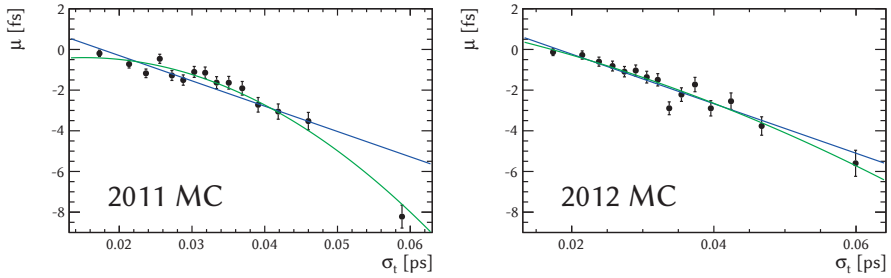


Figure 6.9: μ versus the mean σ_t per bin obtained from a simultaneous fit in bins of σ_t to the $t - t_{\text{true}}$ distribution of samples simulated with 2011 and 2012 conditions contianing a $B_s \to J/\psi\,\varphi$ decay.

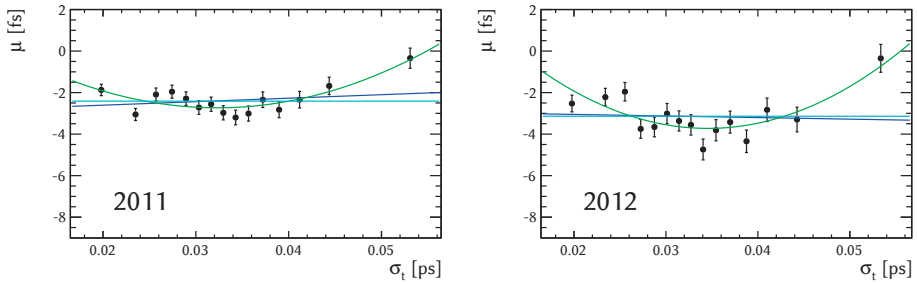
Prompt-Background Samples

To determine whether the common mean of the Gaussians, μ depends on σ_t in the calibration sample and in simulated prompt-background samples, similar simultaneous fits are performed for these samples. The common mean, μ, is again a separate parameter for each bin in σ_t. The result of these fits is shown in figs. 6.10a and 6.10b. As can be seen from the figures, no clear dependence is present and the use of a single value across the bins in σ_t is a valid assumption. This implies that for the prompt-background samples a constant value is sufficient, while for the simulated samples containing a signal $B_s \to J/\psi\,\varphi$ decay, a quadratic parameterisation of $\mu(\sigma_t)$ is required.

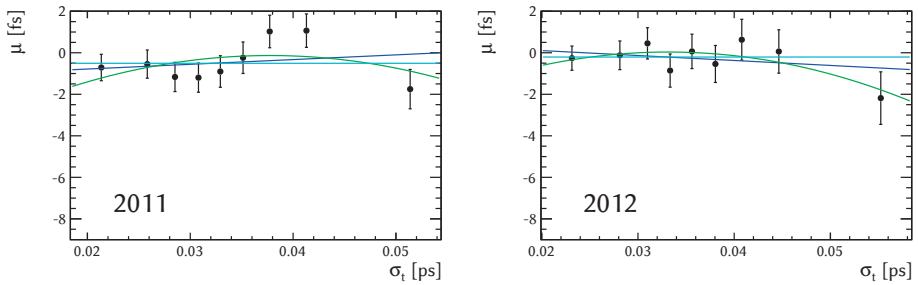
Differences between Simulated Samples

The values of $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, σ_σ: c_0, σ_σ: c_1, μ_0, μ_1 and μ_2 listed in table 6.8 provide the set of calibration parameters for simulated signal samples. These values are compared to those obtained from a sample of simulated prompt-background candidates listed in table 6.7. Differences are observed between the values of the $\overline{\sigma}$: c_1 and σ_σ: c_1 parameters; for simulated $B_s \to J/\psi\,\varphi$ signal candidates their absolute values are larger and have become positive. The reason for this change is not currently understood.

The other significant difference is that the common mean, μ, is not observed to depend on σ_t for prompt-background candidates. The dependence observed in simulated signal decays is therefore concluded to be a feature of signal candidates only. Since there is no calibration sample of signal candidates available in data, a systematic uncertainty must be determined and assigned accordingly.



(a) Data calibration sample.



(b) Simulated sample containing prompt-background candidates.

Figure 6.10: Values of μ versus σ_t obtained from a simultaneous fit. The values estimated by the fit in each bin are shown as the data points. A fit of a constant, a linear and a quadratic polynomial to the points are also shown.

The differences between the resolution model parameters obtained from simulated signal samples and those obtained from simulated prompt-background candidates are combined with the calibration parameters obtained from data samples to yield a set of parameters suitable for estimating the systematic uncertainty; this procedure is described in section 6.3.4.

The origin of the dependence of μ on σ_t for signal candidates is not entirely understood. The fact that it is not present for prompt-background candidates might indicate that the effect depends on the decay-time, since the difference in the decay-time distribution — zero for prompt versus non-zero for signal candidates — is a significant difference between the samples. Since the systematic uncertainties resulting from not including the dependence of μ on σ_t in the final fit are small compared to other systematic uncertainties, assigning a systematic uncertainty will likely remain appropriate for future rounds of this analysis.

6.3.3 Calibration of the Effect of Momentum

From eq. (4.12) it can be seen that the B-momentum uncertainty σ_p only affects candidates with non-zero decay-time. Large effects are not expected and to obtain an indication of the need for additional calibration, figure fig. 6.11 shows the distribution of $t - t_{\text{true}}$ in slices of t_{true} for a simulated sample containing $B_s \to J/\psi\,\varphi$ decays. For each slice in t_{true}, the PDF fitted to the entire sample has been superimposed.

The pull distributions shown at the bottom of each sub-figure in fig. 6.11 are not as expected for the first few bins in t_{true}, which indicates that the PDF does not fit the data very well in these bins. Since the contribution to σ_t of momentum uncertainty is small at low decay-times and the pull distributions for the bins at higher true decay-times look regular, the observed effects are assumed to be unrelated to the contribution of the momentum uncertainty to σ_t. In fact, the irregularities have been identified as caused by the refitting of primary vertices, which is addressed in section 6.4.

6.3.4 Systematic Uncertainties

Two differences are observed between simulated samples of prompt-background and signal decays: differences in the dependence of the width of the Gaussian components of the decay-time resolution model on the reconstructed decay-time uncertainty, σ_t, and differences in the dependence of the common mean of the Gaussians, μ, on σ_t. Since the resolution for real data signal candidates is calibrated on real data prompt-background candidates, such differences lead to systematic uncertainties.

The default set of calibration parameters is obtained from a fit to the prompt-background data sample and called $\mathbf{P}_{\text{D}}$. For simulated samples of prompt background candidates a corresponding set of parameters is extracted, called $\mathbf{P}_{\text{MC}}$. A set of parameters $\mathbf{S}_{\text{MC}}$ is additionally extracted from simulated samples containing $B_s \to J/\psi\,\varphi$ decays. By taking the differences between $\mathbf{P}_{\text{MC}}$ and $\mathbf{S}_{\text{MC}}$, a correction to the data parameters can be formed, resulting in the set of parameters $\mathbf{S}_{\text{D}}$. The set of calibration parameters $\mathbf{S}_{\text{D}}$ is then used in the final fit and the result is compared with the nominal result obtained using $\mathbf{P}_{\text{D}}$. The differences observed for the physics parameters of interest are used as a measure of the systematic uncertainties.

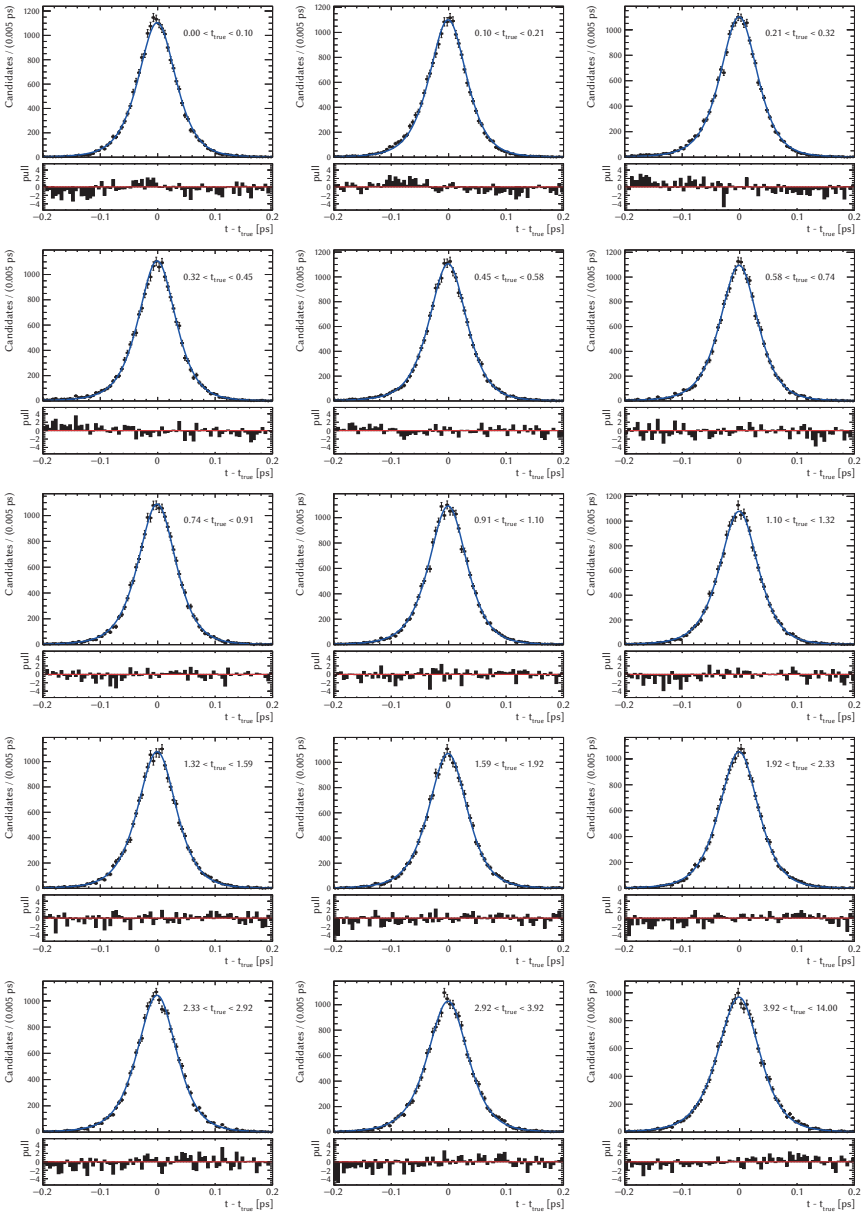


Figure 6.11: Distribution of $t - t_{\text{true}}$ of a sample simulated with 2011 conditions, divided in bins of t_{true}; the curve is the projection of the PDF which was fitted to the entire sample.

Scaling Transformations

When determining the systematic uncertainty related to dependence of the widths of the Gaussian components of the resolution model on σ_t, $\mathbf{P}_{\mathrm{D}}$, $\mathbf{S}_{\mathrm{D}}$, $\mathbf{P}_{\mathrm{MC}}$, and $\mathbf{S}_{\mathrm{MC}}$ are a sets of values of the parameters $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, f_2, σ_σ: c_0 and σ_σ :c_1. To transport the parameters from the prompt-background sample to the signal sample a scaling transformation is assumed to be appropriate in this case.

The scaling transformation for the dependence of the true resolution on σ_t, $\mathbf{T}_{\mathrm{P}\to\mathrm{S}}$, is defined as:

$$\mathbf{T}_{\mathrm{P}\to\mathrm{S}}\mathbf{A}^{-1}\mathbf{P}_{\mathrm{MC}} = \mathbf{A}^{-1}\mathbf{S}_{\mathrm{MC}}, \tag{6.13}$$

where $\mathbf{P}$ is taken from table 6.7, $\mathbf{S}$ is taken from table 6.8 and $\mathbf{A}$ is the matrix of eigenvectors of the prompt parameters' covariance matrix $\mathbf{C}$. By transforming to the eigenbasis of $\mathbf{P}_{\mathrm{MC}}$, the parameters become uncorrelated, which allows simple scaling and ensures that $\mathbf{T}_{\mathrm{P}\to\mathrm{S}}$ is diagonal. From this and eq. (6.13) it follows that:

$$\mathbf{T}_{\mathrm{P}\to\mathrm{S},ii} = \frac{\mathbf{S}'_{\mathrm{MC},i}}{\mathbf{P}'_{\mathrm{MC},i}} \tag{6.14}$$

where $\mathbf{P}' = \mathbf{A}^{-1}\mathbf{P}$ and $\mathbf{S}' = \mathbf{A}^{-1}\mathbf{S}$. $\mathbf{T}_{\mathrm{P}\to\mathrm{S}}$ is then applied to $\mathbf{P}_{\mathrm{D}}$ to obtain $\mathbf{S}_{\mathrm{D}}$:

$$\mathbf{S}_{\mathrm{D}} = \mathbf{A}^{-1}\mathbf{T}_{\mathrm{P}\to\mathrm{S}}\mathbf{A}\mathbf{P}_{\mathrm{D}}. \tag{6.15}$$

The values of $\mathbf{P}_{\mathrm{D}}$ are taken from table 6.5 and the resulting values of $\mathbf{S}_{\mathrm{D}}$ are listed in table 6.9; since the parameters are fixed in the fit, uncertainties have not been propagated and have been omitted from the table.

Additive Transformations

For the determination of the systematic uncertainty related to the dependence of μ on σ_t, $\mathbf{P}_{\mathrm{D}}$, $\mathbf{P}_{\mathrm{MC}}$ and $\mathbf{S}_{\mathrm{MC}}$ consist of the values of μ_0, μ_1 and μ_2. In this case , an additive transformation is assumed to be appropriate to obtain $\mathbf{S}_{\mathrm{D}}$ from $\mathbf{P}_{\mathrm{D}}$:

$$\mathbf{S}_{\mathrm{D}} = \mathbf{P}_{\mathrm{D}} + \left(\mathbf{S}_{\mathrm{MC}} - \mathbf{P}_{\mathrm{MC}}\right), \tag{6.16}$$

where $\mathbf{P}_{\mathrm{D}}$, $\mathbf{S}_{\mathrm{MC}}$ and $\mathbf{P}_{\mathrm{MC}}$ have been taken from tables 6.5 and 6.8 and table 6.7, respectively. The values of $\mathbf{S}_{\mathrm{D}}$ for the dependence of μ on σ_t are listed in table 6.10.

The systematic uncertainties are obtained by using $\mathbf{S}_{\mathrm{D}}$, as listed in tables 6.9 and 6.10 in the final fit and comparing the result to the nominal result obtained using $\mathbf{P}_{\mathrm{D}}$, which are listed in table 6.5. Any differences in the physics observables of interest are taken as an estimate of the systematic uncertainty.

parameter	value 2011		value 2012	
	$\mathbf{S}_\mathrm{D}$	$\mathbf{P}_\mathrm{D}$	$\mathbf{S}_\mathrm{D}$	$\mathbf{P}_\mathrm{D}$
$\overline{\sigma}$: c_0	1.43	1.42	1.44	1.48
$\overline{\sigma}$: c_1	−7.80	−2.85	5.23	−4.96
f_2	0.157	0.244	0.312	0.239
σ_σ: c_0	0.330	0.378	0.278	0.406
σ_σ: c_1	2.76	−1.55	1.85	−2.88

Table 6.9: Resolution model parameters used to estimate the systematic uncertainty resulting from the calibration of the resolution model with prompt-background events.

parameter	value 2011		value 2012	
	$\mathbf{S}_\mathrm{D}$	$\mathbf{P}_\mathrm{D}$	$\mathbf{S}_\mathrm{D}$	$\mathbf{P}_\mathrm{D}$
μ_0	−3.34 [fs]	−2.55 [fs]	−4.53 [fs]	−3.22 [fs]
μ_1	−0.174	-	−0.195	-
μ_2	−6.64 [ps^{-1}]	-	−6.90 [ps^{-1}]	-

Table 6.10: Resolution model parameters used to estimate the systematic uncertainty resulting from the dependence of the common mean of the Gaussians on σ_t observed in the simulated signal samples.

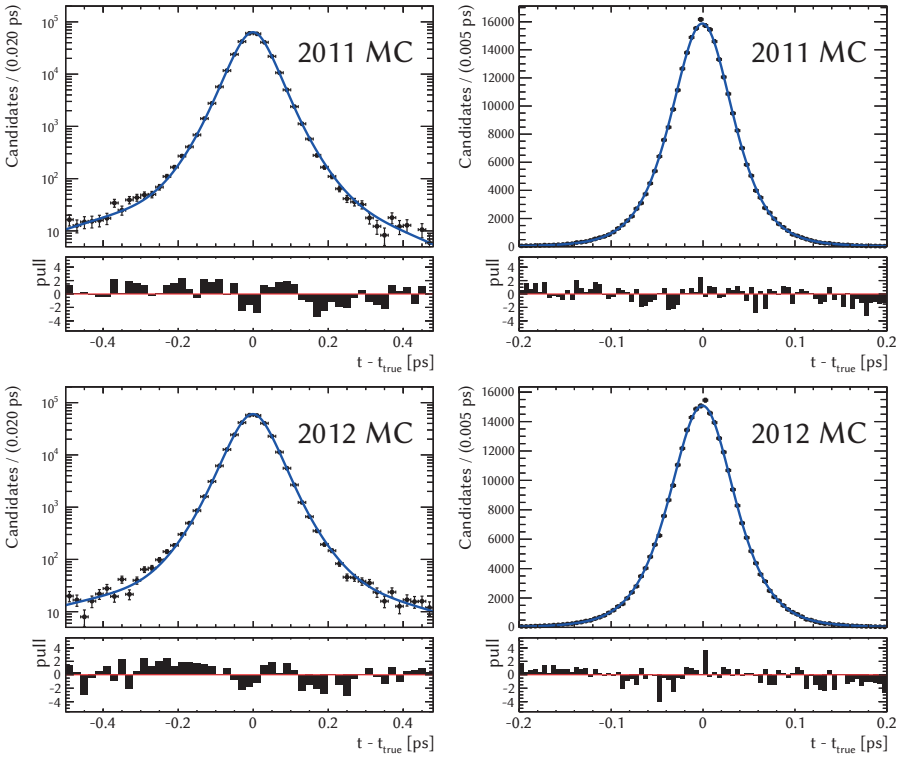


Figure 6.12: Left: result of a likelihood fit to the distribution of $t - t_{\text{true}}$ for samples simulated with 2011 and 2012 conditions containing a $B_s \to J/\psi\varphi$ decay after application of the improved PV refitting. Right: zoomed-in view around $t - t_{\text{true}} = 0$.

6.4 Primary Vertex Refitting

The search for and reconstruction of primary vertices (PVs) is performed once the Velo tracks have been reconstructed and before particle candidates are created. The PV reconstruction therefore has no knowledge of any signal candidates present in the event, and might wrongly include signal tracks of a detached vertex in the primary vertex. As a consequence, this tends to bias the position of the vertex in the direction of the momentum of a signal candidate. An additional effect of the bias is that tracks which lie on the side of the vertex opposite to the direction of the bias might be excluded from the vertex.

Once signal candidates have been reconstructed, the bias due to the inclusion

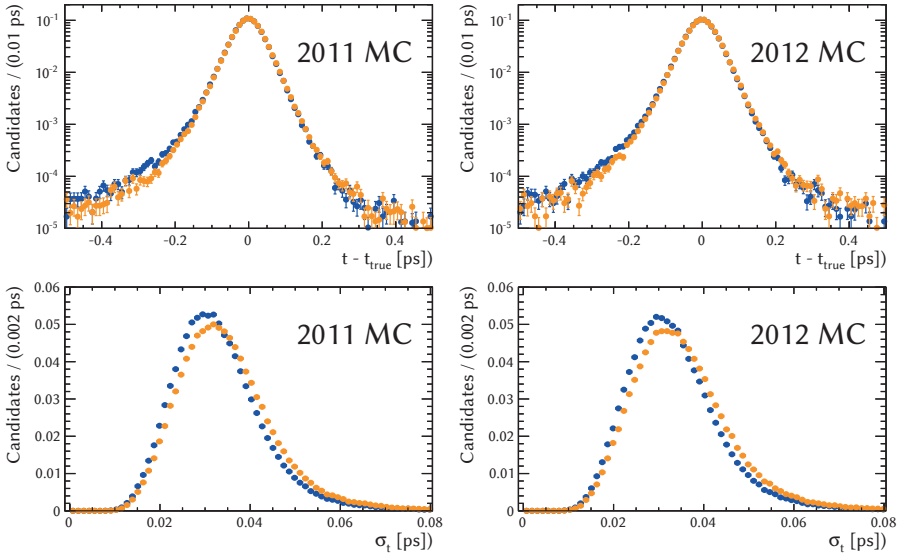


Figure 6.13: Distributions of $t - t_{\text{true}}$ and σ_t with regular (blue) and improved (orange) primary vertex refitting for samples simulated with 2011 and 2012 conditions containing $B_s \to J/\psi\,\varphi$ decays.

parameter	value (2011)			value (2012)		
N_{sig}	291 009.0	$\pm$ 54.6		290 685.0	$\pm$ 54.5	
N_{tail}	994.0	$\pm$ 9.2		1067.0	$\pm$ 8.7	
$\overline{\sigma}$: c_0	1.1964	$\pm$ 0.0020		1.2386	$\pm$ 0.0022	
$\overline{\sigma}$: c_1	0.42	$\pm$ 0.17		−0.19	$\pm$ 0.16	
f_2	0.122	$\pm$ 0.010		0.0928	$\pm$ 0.0083	
σ: c_0	0.2668	$\pm$ 0.0045		0.2773	$\pm$ 0.0052	
σ: c_1	1.97	$\pm$ 0.33		1.67	$\pm$ 0.35	
μ_0	−1.437	$\pm$ 0.082	[fs]	−1.508	$\pm$ 0.084	[fs]
μ_1	−0.0718	$\pm$ 0.0073		−0.0744	$\pm$ 0.0074	

Table 6.11: Results of a fit to simulated samples containing a $B_s \to J/\psi\,\varphi$ decay where the improved PV refitting has been applied. The parameters of the component describing the tails have been omitted.

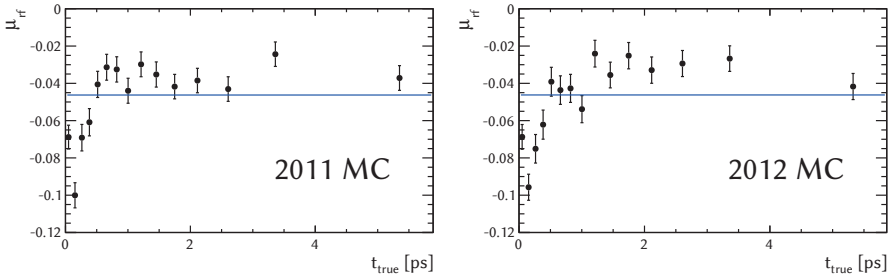
of signal tracks in the PV can be corrected by removing from the PV any tracks which make-up the signal candidate and then recalculating the PV's position. This is the default procedure and removes most biases; it is used everywhere in this thesis, unless stated otherwise. It has the disadvantage that tracks that were excluded from the vertex when it was first reconstructed can never be included again. This effectively results in a residual bias of the position of the PV in the direction of the momentum of signal candidates.

To remove also this residual bias, the complete PV reconstruction algorithm must be run again on all tracks, except those which belong to a given signal candidate. Since the PV reconstruction needs to be repeated for each reconstructed signal candidate, the complete PV refitting procedure is very computationally intensive and therefore only viable for use in the final steps of analyses. In the case of the $B_s \to J/\psi\varphi$ selection procedure, a requirement on each candidates' decay-time is part of the pre-selection, see table 4.2. As a consequence, candidates lost from the selection due to the presence of biases in their decay-time cannot be recovered. The loss of such candidates might result in an additional decay-time dependent acceptance. However, for $B_s \to J/\psi\varphi$, this effect is expected to be small, since $t > 0.2$ ps is required as part of the pre-selection, while a stronger requirement of $t > 0.3$ ps is part of the final selection.

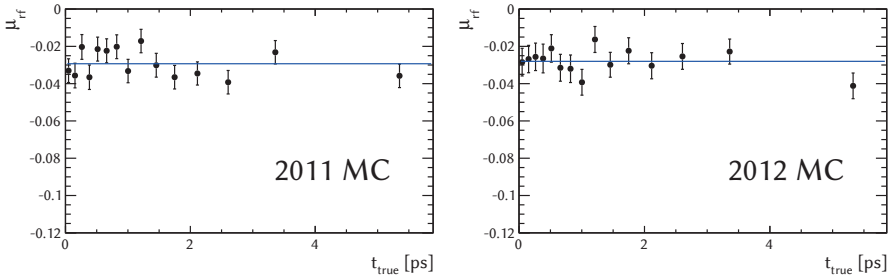
The residual bias in the PV position is partly responsible for the pattern which can be seen in the pulls shown at the bottom of fig. 6.8; once the improved PV refitting was applied the effect became less pronounced, as can be seen by comparing fig. 6.12 to fig. 6.8.

After application of the improved refitting, the asymmetric tails observed in the distribution of $t - t_{\text{true}}$ for simulated signal samples disappear, which has been made explicit in the comparison in section 6.4, which shows the distributions of $t - t_{\text{true}}$ obtained with the default and improved refitting of the primary vertices. The fitted values of $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, σ_σ: c_0 and σ_σ: c_1, listed in table 6.11 are significantly smaller than in the case where the default PV refitting is used, which can be seen by comparing the values with those listed in table 6.8. The reason for this change in the fitted parameters is that the distribution of σ_t has changed, while the distribution of $t - t_{\text{true}}$ has largely remained unchanged, as shown in fig. 6.13.

After refitting, the parameterisation of the common mean of the Gaussian components of the resolution model, μ, as a function of σ_t becomes linear instead of quadratic and the features observed in the pull distributions shown at the bottom of the first four plots of fig. 6.11 disappear. This is concluded from the differences between fig. 6.14a and fig. 6.14b. To obtain these plots, a simultaneous



(a) Sample before improved PV refitting was performed.



(b) Sample where improved PV refitting was performed.

Figure 6.14: Value of the mean of the Gaussian components of eq. (6.9), μ_{rf}, in bins of t obtained from a simultaneous fit to simulated samples containing $B_s \to J/\psi\,\varphi$ decays.

fit in bins of t_{true} of the PDF given by eq. (6.9) was performed to the distribution of $(t - t_{\text{true}})/e\sigma_t$, where μ_{rf} — the mean of the Gaussian components of eq. (6.9) — was determined separately for each bin.

Although the full primary vertex refitting procedure is, as mentioned before, quite computationally intensive, it solves several problems. The distribution of $t - t_{\text{true}}$ becomes more symmetrical, the parameterisation of μ as a function of σ_t becomes linear, μ becomes constant as a function of t_{true} and the estimated decay-time error σ_t becomes closer to the true per candidate resolution. It would therefore be advisable to use the full PV refitting instead of the default PV refitting in a future iteration of the analysis described in this thesis.

6.5 Wrong Primary Vertex Association

To be able to calculate the decay-time of a candidate according to eq. (4.1), it's momentum, decay-vertex and associated origin vertex need to be known. The decay-vertex is obtained by fitting a common vertex to the charged tracks which have been selected as the decay products and the momentum is obtained by adding the reconstructed 4-momenta of the daughter particles. Associating a particle with its correct origin vertex is not straightforward in an environment with pile-up and there are several methods for dealing with this.

In one method, all pairs of candidates and primary vertices are considered and selection cuts are applied to each pair. If more than one pair passes the selection within a single event, one pair can be selected at random or weights can be assigned to the selected pairs. In case weights are assigned, the individual weights should add up to unity.

Another option is to apply a criterion to select the vertex-candidate-pair which is most likely to correspond to the true candidate and its true origin vertex. In this thesis, the vertex is selected with respect to which a candidate has the smallest χ^2_{IP}, defined as the difference of the χ^2 of the PV when including in it or excluding from it the 4-momentum of a candidate.

In both cases a fraction of the selected candidates is associated to a wrong origin vertex and such candidates consequently have a wrongly calculated decay-time. Since such candidates are signal-like in all other respects, they form a dangerous background in the final fit. To treat this background, cuts can either be applied to reduce it to a negligible size, or a component which describes it can be added into the final fit. As described in section 4.1.4, the event selection contains a selection requirement on the second smallest χ^2_{IP} of $\chi^2_{\text{IP}} > 50$ to remove candidates which have likely been wrongly associated to a PV.

To be able to study the consequences of applying this cut and possible improvements arising from removing it, the decay-time distribution of candidates associated to a wrong PV needs to be determined. This has been achieved using an event mixing technique, which is described in more detail in section 6.5.1. The obtained distribution is described in section 6.5.2.

6.5.1 Event Mixing Procedure

To extract from data the decay-time distribution of candidates which are associated to a wrong PV, an event-mixing technique was implemented. Into each event that contains a signal candidate, a vertex from a previous event is artificially injected and flagged as such. The assignment of the best PV is then redone

including the injected vertex and the full selection is applied to the resulting vertex-candidate pairs. Since each injected vertex was flagged, the candidates that pass the final selection and were associated to an injected vertex can be selected. The decay-time distribution of these candidates is then taken as the raw decay-time distribution of candidates with a wrong PV.

Since the injection of a vertex into an event has shifted the distribution of the number of PVs of the sample of mixed candidates after selection, the candidates are assigned weights such that their weighted PV distribution again matches the distribution of the number of PVs in data.

To obtain the resolution model parameters and estimate systematic uncertainties, real data and simulated prompt-background candidates as well as simulated signal candidates are studied. To properly include a component representing candidates with a wrong PV in these three samples, separate decay-time distributions are obtained from their respective event-mixed samples. Once these samples are created, sWeights are used to project-out the required components of the decay-time distribution using the $\mu^+\mu^-$ or the $J/\psi K^+K^-$ invariant mass, as appropriate.

The PDF used to calculate the sWeights for simulated samples containing $B_s \to J/\psi\,\varphi$ decays is defined in eq. (4.3) and the PDF used to calculate sWeights for prompt-background candidates — real data and simulated samples — is given by eq. (6.1). The parameters which that the shape of the PDF used to extract sWeights for the event mixed sample are fixed to the values obtained from a fit to the source sample.

6.5.2 Fits With Event Mixed Distributions

Figure fig. 6.15 shows the decay-time distribution of prompt-background candidates that have been assigned a wrong primary vertex, extracted using the event-mixing procedure described in the previous section. To create a PDF that can be used in a subsequent likelihood fit from the decay-time distribution, the kernel estimation technique described in [86] is used. The PDFs resulting from the application of this technique are shown as the curves in fig. 6.15. From the plots in log scale shown in fig. 6.15, it can be seen that the distribution extends to large decay-times and has an approximately exponential shape above 5 ps. If a significant wrong PV component is present, it enriches the decay-time spectrum at high decay-times. Since candidates with high decay-times give large contributions to the likelihood in the fit of an exponential model, the presence of extra candidates at higher decay-time would bias the measured lifetime, if the wrong PV component is ignored in a fit of the decay-time spectrum.

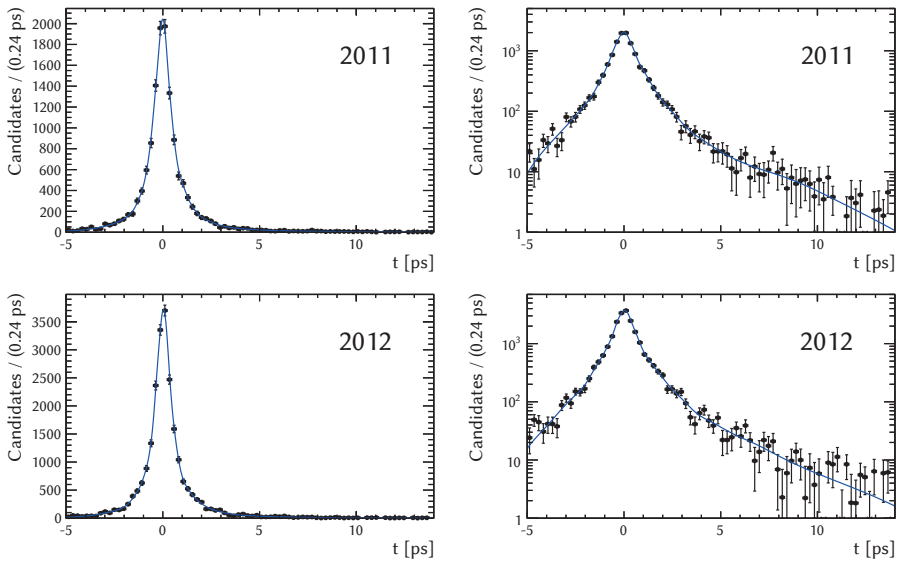


Figure 6.15: Left: decay-time distributions of prompt-background candidates, where sWeights have been calculated with respect to the $\mu^+\mu^-$ invariant mass. Right: the same plots on log scale.

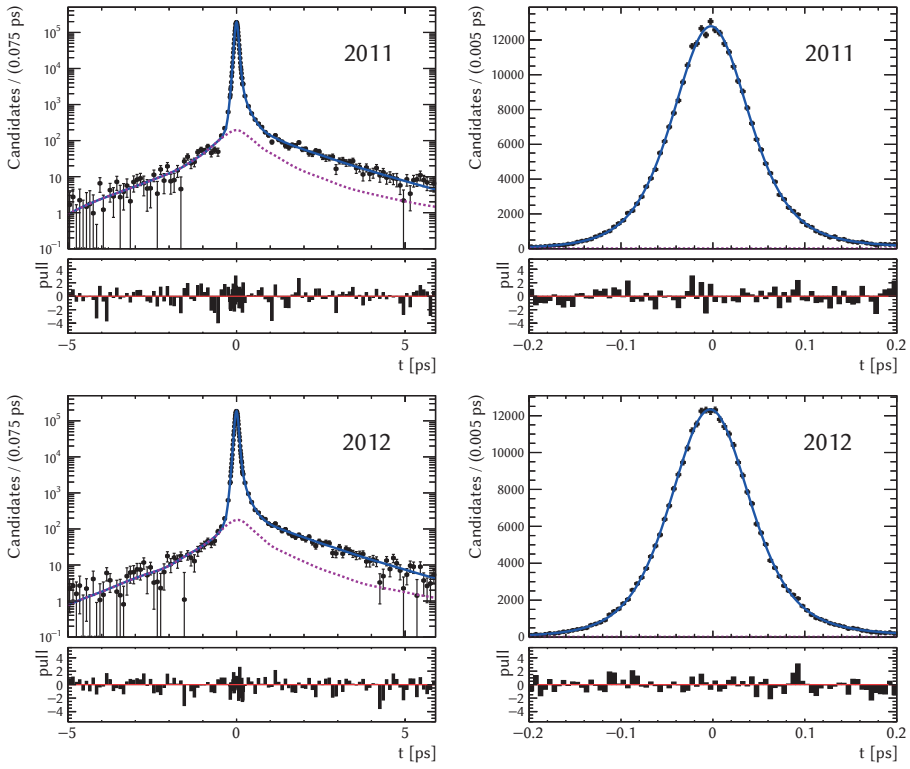


Figure 6.16: Result of a likelihood fit to the 2011 and 2012 real data prompt-background samples, where the next-best-χ^2_{IP} requirement was omitted and a dedicated distribution (dashed purple line) was included to account for candidates which have been assigned a wrong PV.

6.5.3 Data

To test whether the distributions extracted using the event mixing technique indeed describe the decay-time distribution of candidates with a wrong PV in the real data prompt-background sample, a likelihood fit is performed to the decay-time distributions of the 2011 and 2012 real data samples of prompt-background candidates. The fits are performed separately to take the different operating conditions — energy, pile-up, etc. — between 2011 and 2012 into account.

The PDF for these fits is given by eq. (6.2), in which the decay-time distributions of fig. 6.15 replace the wide Gaussian term $\mathrm{G}_{\mathrm{WPV}}(t)$, and the quadratic parameterisation without offset given in eq. (6.6) is used. Since the decay-time

parameter	value 2011	value 2012
N_{prompt}	284 851.0 $\pm$ 62.3	283 573.0 $\pm$ 65.4
N_{ll}	13 893.0 $\pm$ 31.2	15 704.0 $\pm$ 39.2
N_{WPV}	4090.0 $\pm$ 11.4	3512.0 $\pm$ 11.2
f_{sl}	0.7673 $\pm$ 0.0082	0.7752 $\pm$ 0.0074
τ_{sl}	1.464 $\pm$ 0.044 [ps]	1.434 $\pm$ 0.043 [ps]
τ_{ll}	0.1614 $\pm$ 0.0048 [ps]	0.1456 $\pm$ 0.0046 [ps]
$\overline{\sigma}$: c_0	1.4308 $\pm$ 0.0028	1.4840 $\pm$ 0.0030
$\overline{\sigma}$: c_1	$-$0.56 $\pm$ 0.28	$-$2.73 $\pm$ 0.29
f_2	0.251 $\pm$ 0.012	0.256 $\pm$ 0.014
σ_σ: c_0	0.3944 $\pm$ 0.0050	0.4053 $\pm$ 0.0053
σ_σ: c_1	2.43 $\pm$ 0.50	2.90 $\pm$ 0.53
μ	$-$2.421 $\pm$ 0.095 [fs]	$-$3.33 $\pm$ 0.11 [fs]

Table 6.12: Result of a fit to the decay-time distribution of the calibration sample where a dedicated distribution was added to take candidates with a wrong PV into account.

distribution of candidates with a wrong PV is now taken into account correctly, the requirement of $\chi_{\text{IP,next}}(\text{B}_\text{s}) > 50$, designed to remove candidates with a wrong PV, is omitted from the selection.

The results of the fits are listed in table 6.12 and shown in fig. 6.16. By comparing tables table 6.12 and table 6.5 it can be seen that all parameters describing the prompt peak are compatible, except σ_σ: c_1, which has become positive. To investigate the reason for this change, the shape of the wrong PV component is assumed to be independent of decay-time and included in a simultaneous fit to the 2011 real data calibration sample in bins of σ_t, similar to those described in section 6.2.3. The values of $\overline{\sigma}$ and σ_σ in each bin are taken from the fit and shown as the data points in fig. 6.17; the parameterisations defined in eqs. (6.5) and (6.6) are shown as the blue and orange curves, respectively. From fig. 6.17, it can be seen that the linear and quadratic parameterisations both describe the data well, and that the change of sign of σ_σ: c_1 is likely the result of a slight shift upwards of the points.

6.5.4 Simulated Samples

Figure 6.19 shows the result of a fit to the distribution of $t - t_{\text{true}}$ for samples simulated with 2011 and 2012 conditions, where a dedicated distribution was

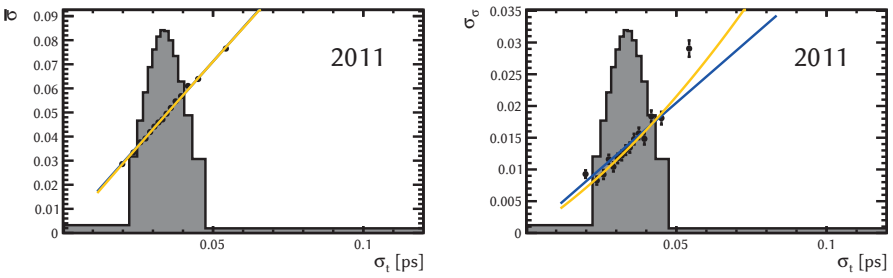


Figure 6.17: Result of a simultaneous fit in bins of σ_t to the 2011 sample of prompt-background candidates. The grey histogram represents the fraction of candidates in each bin. The difference with fig. 6.4 is that the decay-time distribution shown in fig. 6.15 is included instead of a wide Gaussian to take candidates which have been associated to the wrong PV into account.

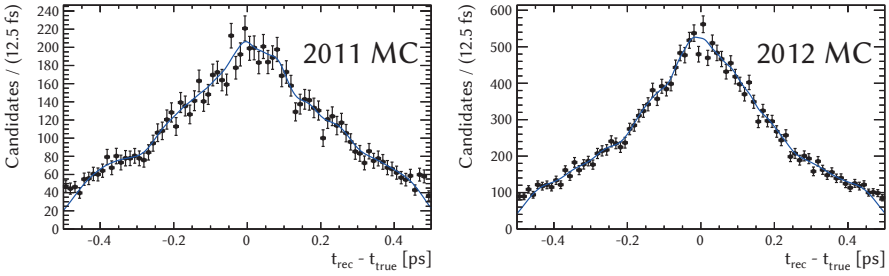


Figure 6.18: Distributions of $t - t_{\text{true}}$ for signal candidates which have been assigned a wrong PV; events simulated with 2011 (left) and 2012 (right) conditions.

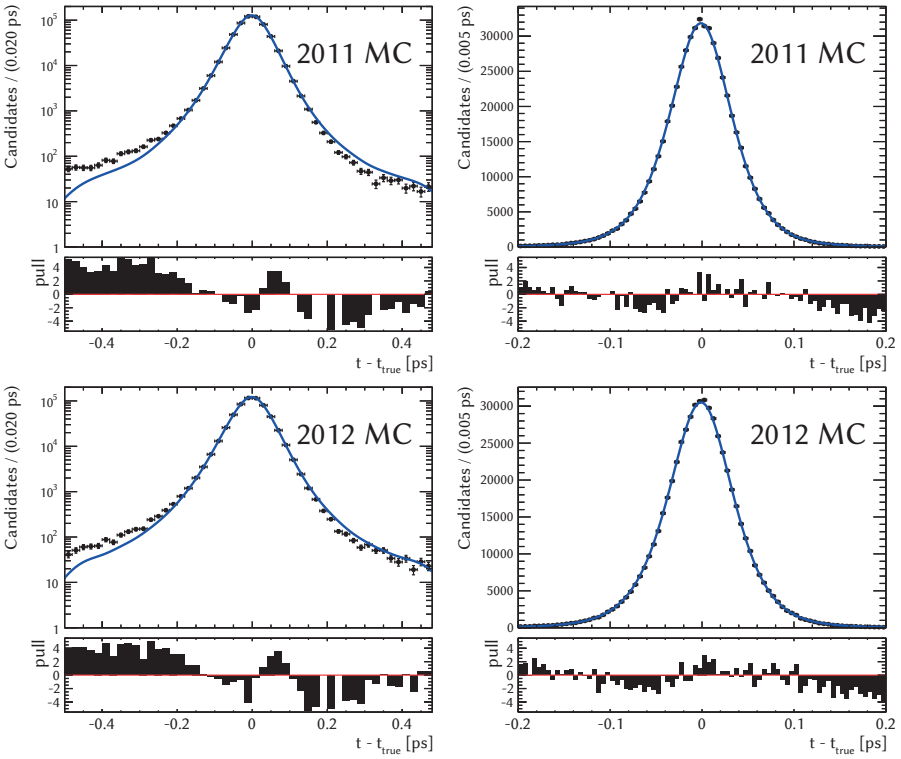


Figure 6.19: Result of a likelihood fit to samples simulated with 2011 and 2012 conditions which contain a $B_s \to J/\psi\,\varphi$ decay, where a dedicated distribution has been included to account for candidates which have been associated to the wrong PV.

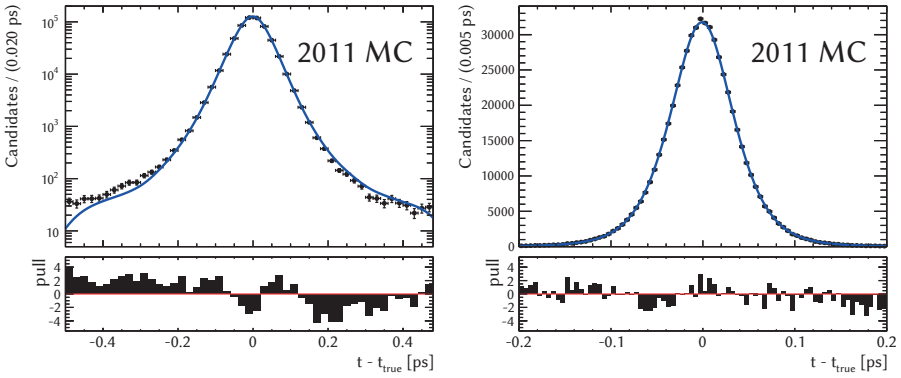


Figure 6.20: Result of a likelihood fit to samples simulated with 2011 conditions which contain a $B_s \to J/\psi\,\varphi$ decay, where a dedicated distribution has been included to account for candidates which have been associated to the wrong PV. The full PV refitting as described in section 6.4 has been applied to the signal candidates, but not during the event mixing procedure.

added to take candidates into account which have been assigned a wrong PV. This distribution has been extracted using the event mixing methods described in the previous section. The correctness of the procedure is evaluated by looking at the residuals after a fit to the distribution of $t - t_{\text{true}}$ of simulated samples containing $B_s \to J/\psi\,\varphi$ decays, where the dedicated distributions for the wrong PV component shown in fig. 6.18 are included. From fig. 6.19 it can be seen that the residuals show some features which indicate a poor fit at $t - t_{\text{true}} < 0$.

Since the improved primary vertex refitting procedure described in section 6.4 was not applied to this sample, some asymmetry is expected in the tails of the distribution of $t - t_{\text{true}}$, as discussed in section 6.4. The distribution of $t - t_{\text{true}}$ for signal candidates associated to a wrong PV is approximately symmetric, as can be seen in fig. 6.18. The result of trying to fit the asymmetric tails in the data distribution with a symmetric PDF is responsible for the poor fit. Figure 6.20 instead shows a fit to a sample where the full PV refitting was applied to the candidates, but not as part of the event mixing procedure. The pulls of the PDF with respect to the data are significantly better than in fig. 6.19, which clearly shows that applying the full PV refitting improves the quality of the fit. It is likely that applying the full PV refitting also as part of the event mixing procedure will further improve the quality of the fit.

For prompt-background samples — real data or simulated — the presence of

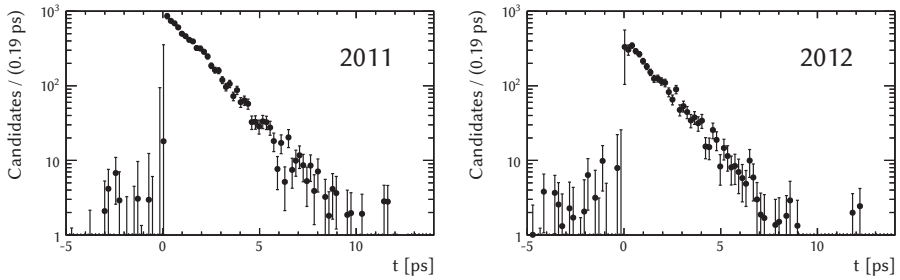


Figure 6.21: Decay-time distribution of signal candidates selected by the prescaled selection.

asymmetrical tails is mitigated by the presence of the component which takes long-lived candidates into account. Since the long-lived component makes the distribution asymmetrical, residual asymmetries — due to the full PV refitting not having been performed — can be absorbed by this component. This might explain why the residual distribution for fig. 6.16 is as expected. It implies that the size of the wrong PV component is likely somewhat overestimated in data.

6.5.5 Signal Candidates with a Wrongly Assigned PV

It has not been possible to directly determine from data the size of the wrong PV component in the signal sample. The components of the PDF that describes the decay-time distribution of signal candidates can only be distinguished from the wrong PV component at large negative decay-times. Due to the low efficiency for signal candidates of the prescaled selection, few signal candidates are available in the sensitive range. This can be judged from fig. 6.21, which shows the decay-time distribution of signal candidates selected by the prescaled selection; sWeights were calculated with respect to the $J/\psi K^+K^-$ invariant mass to obtain the decay-time distribution for signal candidates. A few events are visible at large negative decay-times, but too few to allow a determination of the size of the wrong PV component.

The decay-time for candidates wrongly associated to a PV is obtained from the calibration sample and to be able to use it in a fit to the signal sample, two assumptions are made. These assumptions take two important differences between the calibration sample and the signal sample into account. The first difference is that the selection requirement of $\chi_{\mathrm{IP,next}}(B_s) > 50$, designed to reject candidates with a wrong PV, is not applied to the calibration sample. The second difference

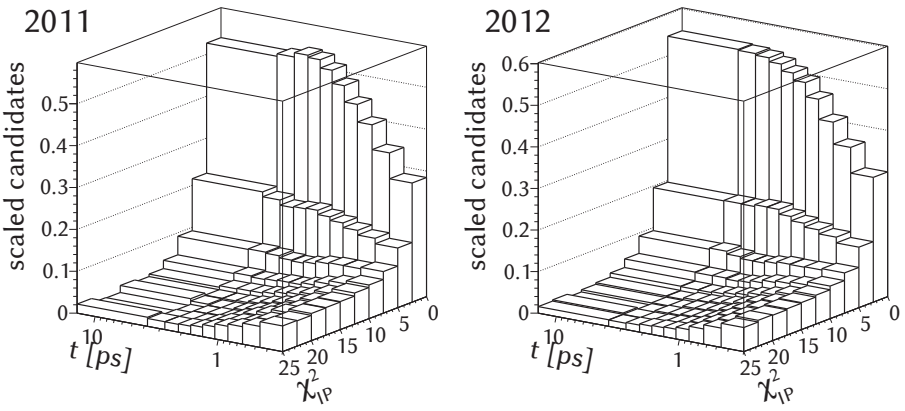


Figure 6.22: Normalised distributions of $\chi^2_{\mathrm{IP}}(\mathrm{B_s})$ in bins of t for 2011 and 2012 signal candidates.

is that the decay-time of the majority of the candidates in the calibration sample is 0 ps while for signal candidates $t > 0.3$ ps.

First, it is assumed that the decay-time distribution of candidates with a wrong PV is independent of the selection requirement of $\chi_{\mathrm{IP,next}}(\mathrm{B_s})$, such that the inclusion of the selection requirement only affects the size of the component, but not its shape. Secondly, it is assumed that the $\chi^2_{\mathrm{IP}}(\mathrm{B_s})$ is independent of its decay-time. This implies that the assignment of a primary vertex to the $\mathrm{B_s} \to \mathrm{J/\psi\varphi}$ candidate is independent of the candidate's decay-time. Figure 6.22 shows normalised distributions of $\chi^2_{\mathrm{IP}}(\mathrm{B_s})$ in bins of t. There are differences between the distributions along the t-axis, in particular they are flatter closer to $t = 0$ and more skewed toward small $\chi^2_{\mathrm{IP}}(\mathrm{B_s})$ at large t, but these are sufficiently small to justify the assumption that $\chi^2_{\mathrm{IP}}(\mathrm{B_s})$ is independent of decay-time.

To extract the size of the wrong PV component, a fit is performed to the 2011 and 2012 prescaled samples which have been selected by the prescaled selection, including the requirement of $\chi_{\mathrm{IP,next}}(\mathrm{B_s}) > 50$, but not the requirement of $t > 0.3$ ps. Using the two assumptions, the fraction of candidates with a wrong PV obtained from this fit becomes valid for the signal sample. From table 6.13 this fraction is calculated to be 0.0025 ± 0.0002 and 0.0016 ± 0.00015 for the 2011 and 2012 samples, respectively. To determine the effect of ignoring this component in the final fit, the fit is performed with this component included, and the differences with the nominal fit result are taken as systematic uncertainties. Since the size of the wrong PV component is likely slightly overestimated in data, the obtained

parameter	value 2011	value 2012
N_{prompt}	290 528.0 $\pm$ 62.1	291 839.0 $\pm$ 64.9
N_{ll}	14 532.0 $\pm$ 32.8	15 467.0 $\pm$ 37.8
N_{WPV}	582.0 $\pm$ 4.6	495.0 $\pm$ 4.7
f_{sl}	0.7510 $\pm$ 0.0072	0.2207 $\pm$ 0.0069
τ_{sl}	1.426 $\pm$ 0.035 [ps]	1.308 $\pm$ 0.034 [ps]
τ_{ll}	0.1546 $\pm$ 0.0046 [ps]	0.1499 $\pm$ 0.0045 [ps]
$\overline{\sigma}\ c_0$	1.4260 $\pm$ 0.0027	1.4871 $\pm$ 0.0029
$\overline{\sigma}\ c_1$	$-$2.85 $\pm$ 0.27	$-$4.96 $\pm$ 0.28
f_2	0.214 $\pm$ 0.011	0.207 $\pm$ 0.010
$\sigma_\sigma\ c_0$	0.3902 $\pm$ 0.0047	0.4201 $\pm$ 0.0051
$\sigma_\sigma\ c_1$	$-$1.59 $\pm$ 0.49	$-$2.75 $\pm$ 0.53
μ	$-$2.520 $\pm$ 0.097 [fs]	$-$3.19 $\pm$ 0.10 [fs]

Table 6.13: Result of a fit to the decay-time distribution of the calibration sample, selected including the additional requirement of $\chi_{\text{IP,next}}(\text{B}_\text{s}) > 50$. A dedicated distribution was added to take candidates with a wrong PV into account.

systematics are conservative.

6.5.6 Future Considerations

In general the results of the fits show that the decay-time distributions for candidates associated to the wrong PV extracted using event mixing are correct. For the LHCb upgrade a much higher pile-up is expected[87] which will result in the wrong PV component becoming more prominent. Since the event mixing techniques described in this section are fully data-driven, they will remain valid for analysis after the LHCb upgrade and will be very useful in extracting the decay-time distribution of the wrong PV component from data and taking it into account correctly in decay-time dependent analyses.

Chapter 7

Results

In this chapter, the results of the analysis of the data sample of $B_s \to J/\psi\varphi$ decays collected by the LHCb experiment in 2011 and 2012 are presented. These results were obtained in close collaboration with ref. [81], within the context of a larger working group inside the LHCb collaboration.The estimates of the physics parameters of interest obtained from the weighted maximum likelihood fit are presented in the section 7.1, together with their statistical uncertainties. The systematic uncertainties are given in section 7.2 and each systematic uncertainty is discussed in more detail in sections 7.2.1 to 7.2.11. The final conclusions are presented in section 7.3.

7.1 Results

The result of the weighted maximum likelihood fit of the PDF given in eq. (4.33) to the combined 2011 and 2012 data sample of selected $B_s \to J/\psi\varphi$ decays is given in table 7.1 and the projections of the PDF on the data as a function of decay-time and angles are shown in fig. 7.1. The -$\Delta\log\mathcal{L}$ is shown in fig. 7.2 as a function of parameter value for ϕ_s and $|\lambda_s|$, and can be seen to closely follow a parabolic shape around the minimum. This indicates that the assumption of Gaussian uncertainty for these parameters is valid.

The weak mixing phase ϕ_s results in a value that is consistent with no *CP* violation to within one sigma. The value of $|\lambda_s|$ shows that the decay $B_s \to J/\psi\varphi$ is dominated by a single diagram. Although no tagged flavour oscillations can be seen in the total decay rate, the parameter Δm_s get its sensitivity from the interference terms in the decay-rate equations. It is consistent with the measurement from the flavour specific decay $B_s \to D_s^- \pi^+$ [88]. The measurement of

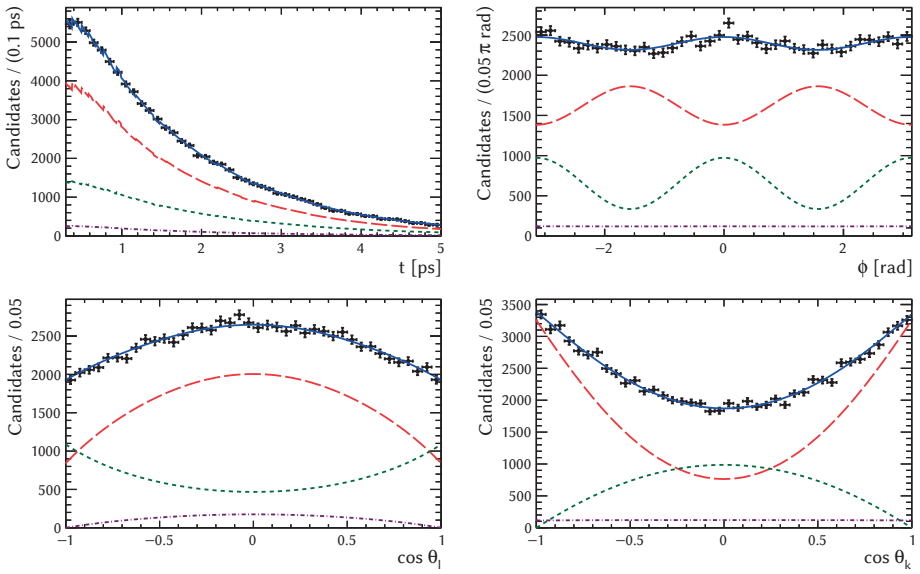


Figure 7.1: Projection of the PDF resulting from the fit to the combined 2011 and 2012 data, superimposed on the data sample as a function of decay-time and decay-angles. The total PDF (blue), *CP*-even (dashed red), *CP*-odd (dashed green) and S-wave (dashed purple) components are shown. Figures produced in collaboration with [81].

the average decay with, Γ_s, is consistent with the dedicated measurement of the lifetimes reported in [79]. The decay width difference, $\Delta\Gamma_s$, is the best measurement with a single channel to date. The magnitudes of the *CP*-odd and *CP*-even polarisation states, $|A_0^{CP}|^2$ and $|A_\perp^{CP}|^2$ are reported, and the magnitude of the third component, $|A_\parallel|^2$, can be obtained from the relation $|A_\parallel|^2+|A_\perp|^2+|A_0|^2 = 1$. The *CP*-even phase of this component is chosen to be equal to zero and the difference between this phase and those of the other polarisations, $\delta_\parallel - \delta_0$ and $\delta_\perp - \delta_0$, are reported. The final set of results are the fraction of the non-resonant (S-wave) decays of $B_s \to J/\psi K^+K^-$ in each bin in $M(K^+K^-)$ and the differences between their *CP*-even phase and that of the "perpendicular" P-wave polarisation state, $\delta_\perp$.

For these phase differences, the likelihoods are far from Gaussian, as can be seen from fig. 7.3. The values of the -$\Delta\log\mathcal{L}$ are close to the minimum value over a much larger range than would justify a "one-sigma" uncertainty to be assigned. For this reason, the interval in which -$\Delta\log\mathcal{L} < 4.5$ is given instead of a central

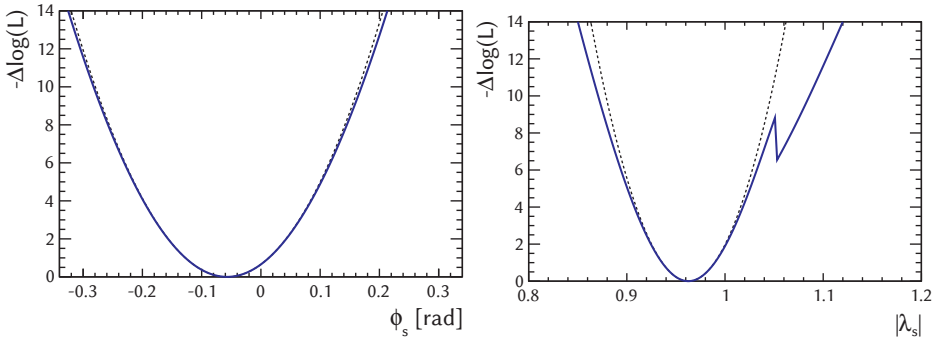


Figure 7.2: Value of -$\Delta\log\mathcal{L}$ as a function of parameter value for ϕ_s and $|\lambda_s|$; ideal, parabolic, likelihoods, computed using the second derivative at the minimum, are shown as the dashed lines. Figures produced in collaboration with [81].

value, which corresponds to a "three-sigma" interval had the uncertainties been Gaussian.

7.2 Systematic Uncertainties

Table 7.2 summarises the systematic uncertainties on all physics parameters of interest. They are discussed in more detail in the following sections. The uncertainties related to the calibration of the flavour tagging are not listed in the table, as they are included in the uncertainty of the final fit through constraints on the calibration parameters. The constraints propagate the total uncertainties on the flavour tagging calibration parameters, which are the quadrature sum of their statistical uncertainties and their systematic uncertainties, to the physics observables. The statistical uncertainties result from the finite size of the calibration samples, while systematic uncertainties arise from the different distribution of kinematic observables between calibration samples and $B_s \to J/\psi\varphi$ decays [80].

To determine the contribution of the constraints, a fit is performed where these parameters are fixed to their nominal value. The central values obtained from this fit are found to be identical to those obtained previously. As a result, the systematic uncertainties arising from the flavour tagging calibration are considered to be negligible. For ϕ_s this becomes clear by applying propagation of uncertainties to the observable decay-rate equations (eq. (4.33)), and noting

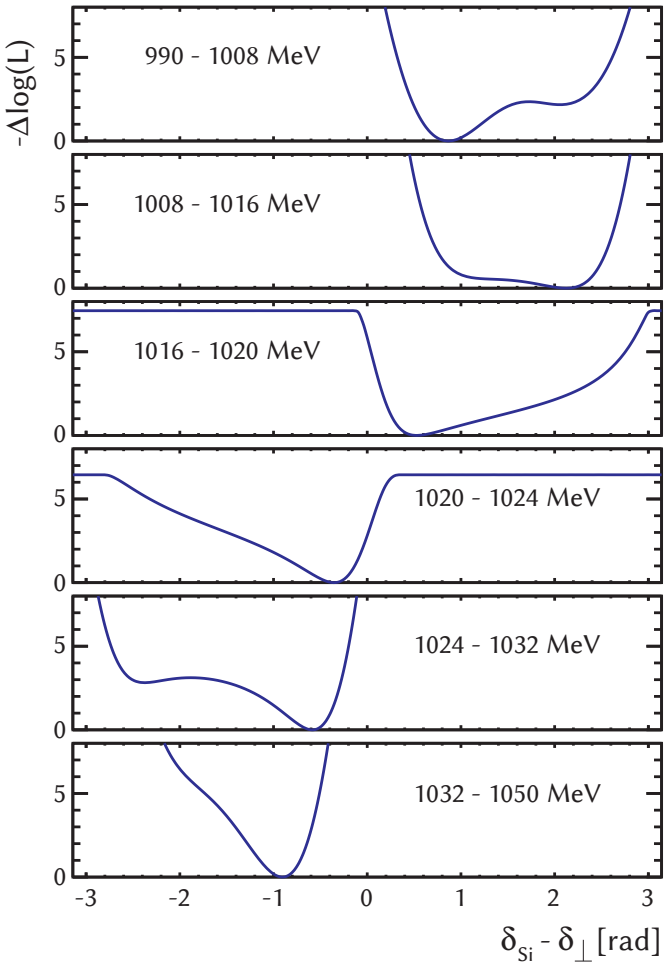


Figure 7.3: Value of -Δlog$\mathcal{L}$ as a function of parameter value for the phase differences between the S-wave component and $\delta_\perp$ in the bins of $M(\mathrm{K}^+\mathrm{K}^-)$. Figures produced in collaboration with [81].

parameter	value	parameter	value
ϕ_s [rad]	-0.057 ± 0.050		
$\lvert\lambda_s\rvert$	0.9627 ± 0.0188		
Γ_s [ps^{-1}]	0.6592 ± 0.0033		
$\Delta\Gamma_s$ [ps^{-1}]	0.0785 ± 0.0092		
Δm_s [ps^{-1}]	17.723 ± 0.057		
$\lvert A_0^{CP}\rvert^2$	0.5237 ± 0.0034	$\delta_\parallel - \delta_0$ [rad]	$3.257^{+0.100}_{-0.172}$
$\lvert A_\perp^{CP}\rvert^2$	0.2512 ± 0.0049	$\delta_\perp - \delta_0$ [rad]	$3.099^{+0.141}_{-0.151}$
F_{S1}	0.426 ± 0.054	$\delta_{S1} - \delta_{\perp 1}$ [rad]	$[0.3, 2.6]$
F_{S2}	0.059 ± 0.018	$\delta_{S2} - \delta_{\perp 2}$ [rad]	$[0.6, 2.7]$
F_{S3}	$0.010^{+0.007}_{-0.006}$	$\delta_{S3} - \delta_{\perp 3}$ [rad]	$[0.1, 2.7]$
F_{S4}	$0.009^{+0.006}_{-0.005}$	$\delta_{S4} - \delta_{\perp 4}$ [rad]	$[-2.3, 0.1]$
F_{S5}	0.048 ± 0.015	$\delta_{S5} - \delta_{\perp 5}$ [rad]	$[-2.7, -0.2]$
F_{S6}	0.192 ± 0.025	$\delta_{S6} - \delta_{\perp 6}$ [rad]	$[-1.6, -0.6]$

Table 7.1: Result of the weighted maximum likelihood of the combined 2011 and 2012 data sample of selected $B_s \to J/\psi\,\varphi$ decays. The parameter $\delta_S - \delta_\perp$ has multiple poorly separated minima in all K^+K^- mass bins, and the interval in which -$\Delta\log\mathcal{L} < 4.5$ is given instead of a central value.

that the uncertainties on the flavour tagging calibration parameters enter the uncertainty on ϕ_s with a multiplication factor of sin ϕ_s, which is small.

7.2.1 Mass Model Uncertainties

When sWeights are calculated to statistically subtract backgrounds, as described in section 4.2, uncertainties on the parameters of the mass model are ignored. To test whether this is valid, a multivariate Gaussian distribution is created from the parameters and covariance estimated by the fit of the mass model to the data sample. A set of parameters is drawn from this distribution five thousand times and a set of sWeights is calculated using each set of parameters. The final fit is then repeated using each of the sets of sWeights and the resulting parameter estimates are compared to the nominal ones; observed differences in the parameter estimated are assigned as systematic uncertainties.

7.2.2 Reflection Backgrounds

Systematic uncertainties related to the subtraction of reflection backgrounds are given the row labelled “reflection backgrounds” in table 7.2. They are the sum in

	ϕ_s [rad]	$\lvert\lambda_s\rvert$	Γ_s [ps^{-1}]	$\Delta\Gamma_s$ [ps^{-1}]	Δm_s [ps^{-1}]	$\lvert A_0^{CP}\rvert^2$	$\lvert A_\perp^{CP}\rvert^2$	$\delta_\parallel - \delta_0$ [rad]	$\delta_\perp - \delta_0$ [rad]
Reflection backgrounds	0.003	0.003	—	0.0004	0.001	0.0002	0.0004	0.024	0.02
Factorisation	0.003	0.001	—	0.0007	0.004	0.0064	0.0031	0.050	0.048
Signal weights (stat.)	—	—	0.0001	0.0008	—	0.0001	—	—	—
K^+K^- mass integrals	—	0.001	—	—	0.002	—	—	0.01	0.01
$B_c^+ \to B_s X$	—	—	0.0005	—	—	—	—	—	—
Decay-time resolution μ	0.002	—	—	—	0.003	—	—	0.001	0.005
Decay-time resolution σ	0.001	0.001	—	—	0.004	—	—	0.003	0.002
Wrong PV assignment	—	0.002	—	—	0.002	—	—	0.004	—
Decay-time acceptance	—	0.001	0.0013	0.0023	0.012	0.0006	0.0005	—	—
Decay angle resolution	0.004	0.001	—	0.0008	0.006	0.0001	0.0006	0.036	0.022
Angular acceptance (stat.)	0.004	0.002	—	0.0002	0.001	0.0004	0.0011	0.019	0.009
Angular acceptance (sim.)	0.001	0.005	—	—	0.002	0.0020	0.0011	0.006	0.002
total systematic	0.007	0.007	0.0014	0.0027	0.015	0.0067	0.0036	0.069	0.058
statistical	0.051	0.0188	0.0033	0.0092	0.060	0.0034	0.0049	0.132	0.165
total	0.05	0.020	0.004	0.010	0.06	0.008	0.006	0.15	0.17

Table 7.2: Systematic uncertainties for physics parameters of interest.

quadrature of the systematic uncertainties resulting from the uncertainty on the yields of the reflection backgrounds and of the possible differences between data and simulated samples of $B^0 \to J/\psi K^*$ and $\Lambda_b \to J/\psi p K$ decays. The former are determined by varying the subtracted yields by one standard deviation, repeating the final fit, and assigning observed differences in the parameter estimates as systematic uncertainties.

Systematic uncertainties related to the differences between data and simulated samples of $B^0 \to J/\psi K^*$ and $\Lambda_b \to J/\psi p K$ decays are evaluated by reweighing the simulated $B^0 \to J/\psi K^*$ and $\Lambda_b \to J/\psi p K$ samples to reflect the properties of these decays observed in data, and subtracting these reweighed samples before repeating the final fit. Parameter estimates obtained from this fit are compared to the nominal results — obtained without reweighing prior to subtraction of the $B^0 \to J/\psi K^*$ and $\Lambda_b \to J/\psi p K$ backgrounds — and observed differences are assigned as systematic uncertainties.

7.2.3 Factorisation

To be able to use sWeights to statistically subtract backgrounds, the assumption that $\mathrm{PDF}(m, t, \boldsymbol{\Omega}) = \mathrm{PDF}(m) \times \mathrm{PDF}(t, \boldsymbol{\Omega})$ is made. To test this assumption, the data sample was sub-divided into four sub-samples in the decay-time and decay-angles separately, and the invariant mass fit and subsequent calculation of sWeights were performed on each of the four resulting samples, followed by the final fit. Differences in the parameter estimates for the physics parameters of interest were only observed for the binning in $\cos\theta_\mu$ and assigned as systematic uncertainties, which are given in the row labelled "factorisation" in table 7.2. These systematic uncertainties are a significant contribution to the total systematic uncertainty on most parameters, and the dominant contribution for $|A_0^{CP}|^2$, $|A_\perp^{CP}|^2$, $\delta_\parallel - \delta_0$, $\delta_\perp - \delta_0$. To reduce this systematic uncertainty in the future, the background subtraction could be performed in bins of $\cos\theta_\mu$.

7.2.4 K^+K^- mass integrals

To avoid dependence of the observable PDF on the K^+K^- invariant mass, the final fit is performed simultaneously in six bins in the K^+K^- invariant mass. As described in section 4.3, integration of the PDF in these bins results in the appearance of so-called C_{SP}-factors in the "binned" PDF. The C_{SP}-factors are obtained by calculating an integral of the product of the P- and S-wave line shapes over the bin. Differences between data simulated events and the limited knowledge of the S-wave line shape lead to systematic uncertainties. These two contribution

are evaluated separately and added in quadrature to give the systematic uncertainties given in the row labelled "K^+K^- mass integrals" in table 7.2.

The effect of differences between data and simulated events on the invariant K^+K^- mass are taken into account via an alternative set of C_{SP}-factors, calculated under the assumption of a 20% worse resolution. Observed differences between nominal result and the parameter estimates from this fit are included as systematic uncertainties.

The systematic uncertainty for the limited knowledge of the S-wave line shape is obtained by recalculating the C_{SP}-factors under the assumption of a uniform (constant) line shape, and repeating the final fit using the alternative C_{SP}-factors. The resulting differences in the parameter estimates between the nominal fit and this alternative fit are included as systematic uncertainties.

7.2.5 $B_c^+ \to B_s X$

The decay $B_c^+ \to B_s\pi^+$ was recently observed [89], and from this measurement it is estimated that 0.8% of the selected $B_s \to J/\psi\,\varphi$ decays originated from a $B_c^+ \to B_s\pi^+$ decay [80]. When the decay-time of the B_s is calculated, it is assumed to originate from the assigned primary vertex, which is not correct in this case. The measured decay-time is, therefore, overestimated, and a systematic uncertainty is assigned for ignoring this in the final fit.

This systematic uncertainty is evaluated by performing 350 simulated experiments, where a fit is performed to a sample of simulated and selected $B_s \to J/\psi\,\varphi$ decays of equal size to the data sample, including an 0.8% contribution of decays where the B_s originates from a $B_c^+ \to B_s\pi^+$ decay. The $B_c^+ \to B_s\pi^+$ decays are randomly sampled from the sample of decays that passes the $B_s \to J/\psi\,\varphi$ selection. A bias on Γ_s equal to 0.2σ is observed, which is added as a systematic uncertainty; all other parameters are unaffected.

7.2.6 Decay-Time Resolution

Systematic uncertainties that arise from the calibration of the decay-time resolution model are a result of the fact that a sample of prompt-background candidates is used to calibrate the model instead of a sample of signal candidates. Systematic uncertainties are evaluated using simulated events by determining the transformation of the resolution model parameters obtained from a simulated sample of prompt background parameters to the parameters obtained from a sample of simulated events containing $B_s \to J/\psi\,\varphi$ decays. The transformation is subsequently applied to the parameters obtained from the data sample of

prompt background candidates and the resulting parameters are used in a fit of the selected $B_s \to J/\psi\varphi$ data sample. The difference between the result obtained with this fit and the result listed in table 7.1 is taken as systematic uncertainty.

Two differences are observed between the simulated sample of prompt background candidates and the simulated sample of $B_s \to J/\psi\varphi$ candidates: differences in the parameterisation of the common mean and in the widths of the Gaussian components of the resolution model as a function of σ_t. To evaluate the systematic uncertainties arising from the calibration of both components of the resolution model, the parameters listed in tables 6.9 and 6.10 are used in the final fit. The resulting systematic uncertainties are given in the rows labelled "decay-time resolution μ" and "decay-time resolution σ" in table 7.2, respectively.

7.2.7 Wrong Primary Vertex Assignment

Statistically, a fraction of the selected $B_s \to J/\psi\varphi$ candidates must be assigned to the wrong primary vertex, and thus must have a wrongly reconstructed decay-time. As described in section 6.5.5, the expected fraction of $B_s \to J/\psi\varphi$ candidates with a wrongly assigned primary vertex is estimated to be $0.25 \pm 0.02\%$ and $0.16 \pm 0.015\%$ for the 2011 and 2012 samples, respectively. To estimate the systematic uncertainties arising from neglecting this component in the final fit, a PDF that describes this component is constructed and included in the final fit with a yield constrained using the observed fractions for the 2011 and 2012 data samples. The differences in the parameter estimates between this fit and the nominal fit are taken are assigned as the systematic uncertainties; they are listed in the row labelled "wrong PV assignment" in table 7.2. Since the size of the wrong PV component is likely to be slightly overestimated in data, the obtained systematics are conservative.

7.2.8 Decay-Time Acceptance

Systematic uncertainties related to the modelling of decay-time dependent acceptances are given in the row labelled "decay-time acceptance" in table table 7.2, and consist of two contributions that are added in quadrature: a low-time acceptance turn-on and a high-time acceptance reduction. Both result from the Velo reconstruction. Since the lower decay-time acceptance is determined as part of the likelihood fit, any uncertainties resulting from it are propagated to the physics parameters, and no additional systematic uncertainties are required.

The high decay-time systematic uncertainty related to the observed Velo reconstruction efficiency, described in section 5.1, is split in two parts. The first

part is due to the finite size of the $B^+ \to J/\psi K^+$ sample used to determine the parameters of eq. (5.2). To propagate the statistical uncertainties on the parameters listed in fig. 5.1 to the uncertainty on the physics parameters of interest, the parameters of eq. (5.2) are varied by $\pm\sigma$ when generating the efficiency weights. The observed differences in the parameter estimates are assigned as systematic uncertainties.

The second systematic uncertainty is a result of the assumption that the high decay-time Velo reconstruction efficiency only affects Γ_s. To test whether this assumption is correct, the final fit is repeated including the Velo reconstruction efficiency weights. Differences between parameter estimates resulting from this fit and the nominal fit are taken as systematic uncertainties. This is the dominant contribution to the systematic uncertainties arising from the modelling of the decay-time acceptance, which, in turn, is the largest systematic uncertainty for Γ_s, $\Delta\Gamma_s$ and Δm_s. For future iterations of this analysis, this could be improved by including the Velo reconstruction efficiency weights in the final fit by default.

7.2.9 Angular Resolution

The resolution on the reconstructed decay angles was found to be approximately 10 mrad in simulation. It is ignored in the final fit, which leads to systematic uncertainties. Pseudo-experiments were performed to test the effect ignoring the angular resolution on the parameter estimates and their statistical uncertainties [69]. In these experiments, a data sample is generated and the decay-angles are smeared according to the resolution found in simulation. The final fit is then performed on the smeared and on the unsmeared data sample and the differences between the resulting parameter estimates are saved. The distributions of these differences are divided by the uncertainty estimated by the fit to obtain the so-called pull distribution for each of the physics parameters of interest.

The mean and width of each pull distribution are multiplied by the respective parameter's statistical uncertainty obtained from the final fit to data, and added in quadrature to obtain the systematic uncertainties. They are given in the row labelled "decay angle resolution" in table 7.2.

7.2.10 Angular Acceptance

The angular acceptance, as observed in section 4.6, is determined from simulated events containing $B_s \to J/\psi\,\varphi$ decays, and the size of the simulated sample and differences between data and the simulated samples lead to systematic uncertainties.

To evaluate systematic uncertainties related to the size of the simulated sample of events containing $B_s \to J/\psi\,\varphi$ decays, a multi-dimensional Gaussian distribution is created from the angular acceptance coefficients, their uncertainties and their correlations. A set of coefficients is drawn from this distribution 200 times, and the final fit is repeated using each set. For each physics parameter of interest, the RMS of the distribution of differences between the resulting parameter estimates and the nominal parameter estimate is assigned as a systematic uncertainty, which is given in the row labelled "angular acceptance: statistical" in table 7.2.

Differences between data and simulated events containing $B_s \to J/\psi\,\varphi$ events are accounted for by reweighing the simulated sample to match data, as described in section 4.6.1. Systematic uncertainties for the reweighing are obtained by comparing the nominal parameter estimates to those obtained from the final fit when simulated events are not reweighed; observed differences are conservatively assigned as systematic uncertainties, and are given in the row labelled "angular acceptance: simulation" in table 7.2.

7.2.11 Correlations between Angular and Decay-Time Acceptances

Correlations between the angular acceptance and the decay-time acceptance are assumed to be absent and the two acceptances are assumed to factorise. This was verified by calculating the coefficients c_{lm}^{i} of the angular acceptance separately in bins of decay-time and comparing their values. No significant differences between the sets of parameters was observed.

7.2.12 Discussion

It is important to note that for all physics parameters listed in table 7.1, except $|A_0^{CP}|^2$, the total systematic uncertainty is smaller than the statistical uncertainty; for ϕ_s and $|\lambda_s|$ it is even an order of magnitude smaller. The measurement of these parameters is therefore likely to remain statistically limited during LHC Run II in 2015–2017 and for a significant part of the LHCb upgrade [87].

In this thesis, decay-time acceptance and resolution have been modelled in detail. The acceptance was studied both in the HLT and for the final selection. The construction of an HLT selection that has an efficiency of approximately 70% without including selection requirements on quantities that are correlated with the decay-time proved to be a crucial component in understanding the decay-time resolution and acceptance, which in turn led to small respective systematic uncertainties.

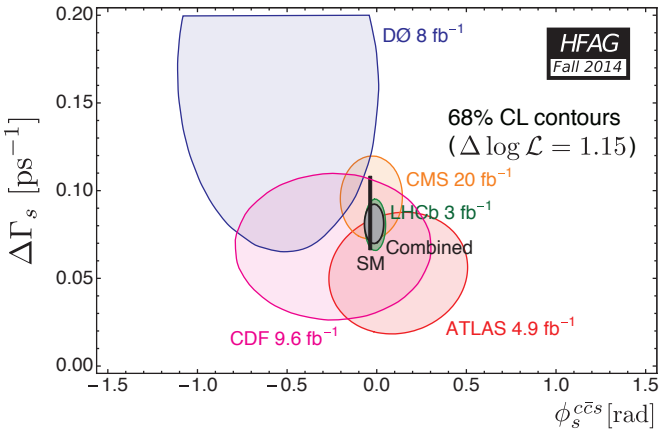


Figure 7.4: Two dimensional contours of -$\Delta\log\mathcal{L}$ =1.15 ("one-sigma") of $\Delta\Gamma_s$ and ϕ_s for all measurements of these observables performed to date. The Standard Model prediction [13, 90], and the combination of all measurements are also shown. Combination performed and figure provided by the Heavy Flavour Averaging Group [91].

In addition, the decay-time distribution of candidates with a wrongly assigned PV was modelled in detail, and found to extend to high decay-times. While for the data sample used in this analysis, this component was found to be sufficiently small to result in small contributions to the systematic uncertainties, it is expected to become more problematic with the increased amounts of pile-up expected for the LHCb upgrade.

7.3 Conclusions

The goal of the analysis described in this thesis is to measure *CP*-violation in the interference between mixing and decay in decays of $B_s \to J/\psi\,\varphi$. The results for ϕ_s, $|\lambda_s|$, Γ_s and $\Delta\Gamma_s$ are:

$$\begin{aligned}
\phi_s &= \text{-}0.057 \quad \pm 0.051 \quad \pm 0.007 \quad \text{rad}, \\
|\lambda_s| &= \quad 0.9627 \pm 0.0188 \pm 0.007 \\
\Gamma_s &= \quad 0.6592 \pm 0.0033 \pm 0.0014 \quad \text{ps}^{-1}, \\
\Delta\Gamma_s &= \quad 0.0785 \pm 0.0092 \pm 0.0027 \quad \text{ps}^{-1},
\end{aligned}$$

where the first uncertainty is statistical and the second is systematic. These are the most precise measurements of these quantities with a single decay channel to date. This also becomes clear from fig. 7.4, which shows the two-dimensional likelihood contours for $\Delta\Gamma_s$ and ϕ_s for the highest precision analysis performed by the CDF [92], DØ[93], ATLAS [94], CMS [95] collaborations, in addition to the LHCb analysis [80]; the LHCb result shown in fig. 7.4 is the combination of the analyses of $B_s \to J/\psi K^+K^-$ and $B_s \to J/\psi \pi^+\pi^-$ decays, as reported in [80]. The combined result of all analysis is also shown, and it can be seen that the LHCb result provides the most precise contribution.

The results for ϕ_s and $|\lambda_s|$ are compatible with no *CP*-violation in $B_s \to J/\psi \varphi$ decays. They are also compatible with the respective Standard Model predictions of $\phi_s = -2\beta_s = -0.0363 + 0.0014 - 0.0012$ rad [13], and $\Delta\Gamma_s = 0.087 \pm 0.021$ ps^{-1} [90], and thus there is no evidence for physics 'beyond the Standard Model'.

In the 2015–2017 run at a beam energies of 13–14 TeV, the statistical uncertainties for this measurement is expected to halve, which would mean that the predicted Standard Model value will still be out of range. For the LHCb upgrade, the error is expected to reduce to a level that will allow testing of the Standard Model prediction. At that stage, it will be important to control the systematic uncertainties, in particular those that result from differences between data and simulated events, such as the angular and decay-time acceptances, and those that are a result of reflection backgrounds. In addition, wrong primary vertex assignment will likely become more problematic, as the pile-up is expected to increase by a factor 2–5 during the LHCb upgrade. Given the current size of the systematic uncertainties resulting from the modelling of the decay-time resolution, they are expected to be under control until the later phase of the LHCb upgrade.

Appendix A

Streamer Framework

For some time before the end of 2010, it had become clear that the implementation of HLT selections needed a serious overhaul for technical reasons, and improved clarity and transparency. The HLT1 process of incremental reconstruction steps (upgrades) of seed tracks is interleaved with selection requirements on the available quantities. It assumed that every objects was a track. Objects such as L0 stubs or calorimeter clusters were converted to tracks by injecting the relevant information into the extra information storage available for a track, which is not intended for this purpose.

The implementation could not easily be improved and an alternative was designed based on a wrapper object which could be transparently passed between upgrades and from which the appropriate underlying objects could be extracted for reconstruction or testing against selection requirements. To interfere as little as possible with the existing code that performs the reconstruction steps, wrapper objects were created to provide the extraction of the correct type of object, memory management and further bookkeeping functionality such as ensuring no CPU-intensive operations are performed twice on the same seed.

To improve the clarity of what an HLT selection is doing, and reuse existing functionality as much as possible the new implementation is based on LoKi functors and streamers [96]. An example of a streamer in HLT1 is implementation of the high mass dimuon selection described in section 3.2.7.

```
VeloCandidates
>>  MatchVeloMuon
>>  LooseForward
>>  ( ( TrPT > 500.0 * MeV ) & ( TrP  > 3000.0  * MeV ) )
>>  IsMuon
```

```
>>  FitTrack
>>  ( TrCHI2PDOF < 3.0 )
>>  MakeDiMuons
>>  ( RV_MASS ( 'mu+' , 'mu-' ) > 2700.0 * MeV )
>>  SINK( 'Hlt1DiMuonHighMassDecision' )
>> ~TC_EMPTY
```

The `VeloCandidates` step runs the Velo raw event decoding and the Velo pattern recognition if they had not been run before and creates wrapper objects from the reconstructed Velo segments. `MatchVeloMuon` applies the Velo segment to muon hit matching algorithm described in section 3.2.2 to the Velo segments, followed by the forward pattern recognition described in section 3.2.4. Selection requirements on total and transverse momentum are imposed followed by the offline muon identification described in section 3.2.5. Identified tracks are fitted and selected based on their χ^2/DoF. The `MakeDiMuons` step creates two-track combinations and selects those with sufficiently small DOCA and χ^2/DoF of the vertex. An invariant mass requirement is imposed on the candidates and accepted candidates are stored by the `SINK`. The selection accepts the events if at least one candidate is stored.

An important role of the functors that implement the upgrade steps is to prevent the same upgrade from being performed twice on identical seeds. The wrapper objects contain a linear history of created objects, that cannot properly accommodate the multiple, branching, paths a seed track might take through the HLT selections. The decision was therefore made to create a new set of wrapper objects for every HLT1 selection as part of the fist step. The underlying objects, tracks in most cases, are the same for all selections. Once an upgrade step is complete, the number of resulting tracks is stored on the seed track with a unique identifier. To prevent failed upgraded from being repeated, 0 is stored on failure. If a seed that was already upgraded by another selection enters the same step again, the number of previously created tracks is retrieved and if there are at least one, the previously created tracks are retrieved from memory and added to the output.

The streamers allow for more operations than passing a set of wrapper objects to the next step by using the streaming operator, `>>`. Conditional statements are possible using the multiplication operator `*`; this operator is used, for example, when the intended upgrade depends on the success of an external algorithm, which is executed on-demand. A tee adaptor is also present, that allows copying of the set of wrapper objects so they can be monitored or passed to a parallel set of upgrade steps.

Appendix B

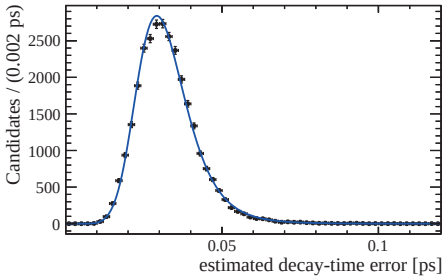
Numerical Calculation of the Dilution

To test the algorithm used to numerically calculate the discreet Fourier transform, toys have been used. A model as given by eq. (4.13) with $N = 2$ has been used to generate a decay-time residual distribution. To generate the per-event estimated decay-time errors, σ_t, a double Log-Normal distribution has been fit to the σ_t distribution of the 2011 signal sample. The parameters obtained from the fit are then used to generate the fast simulations. Figure B.1a shows the distribution of σ_t for this sample and the result of the fit. The Log-Normal distribution is given by:

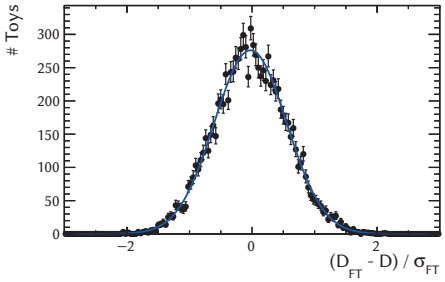
$$P(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2} \tag{B.1}$$

In data the calibration sample is assumed to consist of two components, a combinatorial background component and a component which peaks in the $\mu^+\mu^-$ invariant mass around the J/ψ mass, and sWeights are used to statistically subtract the combinatorial background component. The use of weights in the calculation of the dilution has therefore been included in the validation.

For each toy, distributions of the $\mu^+\mu^-$ invariant mass, decay-time and σ_t are generated, where the parameters of the $\mu^+\mu^-$ invariant mass PDF have been obtained from a fit to the decay-time resolution model calibration sample collected in 2012. After generation, a fit to the $\mu^+\mu^-$ invariant mass is performed, followed by calculation of the sWeights. The dilution is then calculated using both the generated σ_t distribution and eq. (4.16) or directly from the generated decay-time distribution using the Fourier transform method. The distribution of the



(a) Fit of a double Log-Normal PDF to the σ_t distribution of the 2011 signal sample.



(b) Distribution of the difference between the dilution calculated using either the Fourier transform algorithm or the parameters of the PDF used to generate the toy samples.

difference between these two methods divided by the uncertainty obtained from the Fourier transform method is shown on the left side of fig. B.1b; ten thousand toys have been generated with a hundred thousand events per toy. A Gaussian PDF has been fitted to the resulting distribution and the fit yields a mean value of -0.025 ± 0.006 and a width of 0.578 ± 0.004. This shows that the uncertainty is underestimated by a factor 1.73 and that there is a small bias of 2.5% of the uncertainty, which is considered to be negligible.

Appendix C

Effect of Using sWeights for Resolution Fits

This appendix describes the result of a study using pseudo-experiments to determine the effect of the use sWeights on the uncertainties obtained from the weighted likelihood fit to the calibration sample in the context of the calibration of decay-time resolution model on data, as described in chapter 6. The following is a summary of the sWeigting procedure described in this chapter, which is included for clarity. The results of the study are then presented and discussed.

C.1 Use of sWeights for Calibration of the Resolution Model

The calibration of the decay-time resolution-model requires a sample of prompt background candidates that include a $J/\psi \to \mu^+\mu^-$ decay, as described in section 6.1. The sample of prompt-background candidates contains B_s candidates created from two random muons and two random kaons in addition to those created from a J/ψ and two random kaons. Since the kinematic properties of combinatorial background candidates are less similar to signal $B_s \to J/\psi\,\varphi$ candidates, combinatorial background candidates are not used to calibrate the decay-time resolution model. The decay-time distribution of the combinatorial background component is statistically subtracted from the full sample by fitting the $\mu^+\mu^-$ invariant-mass distribution and using the result of the fit to calculate sWeights, as described in section 4.2. The PDF used to fit this sample is given by:

$$\mathrm{PDF}(M) = N_{J/\psi} \times ((f_{\mathrm{CB}}\mathrm{CB}_1(M) + (1 - f_{\mathrm{CB}})\mathrm{CB}_2(M)) + N_{\mathrm{bkg}} \times \mathrm{e}^{-M/c}, \quad \text{(C.1)}$$

parameter	μ	σ	μ cor.	σ cor.
$\bar{\sigma} : c_0$	-0.022 ± 0.022	1.464 ± 0.017	-0.017 ± 0.016	1.019 ± 0.012
$\bar{\sigma} : c_1$	0.009 ± 0.021	1.341 ± 0.018	0.024 ± 0.016	0.996 ± 0.014
f_2	0.091 ± 0.019	1.195 ± 0.039	0.033 ± 0.015	0.941 ± 0.025
$\sigma_\sigma : c_0$	0.007 ± 0.021	1.393 ± 0.021	0.004 ± 0.015	0.952 ± 0.035
$\sigma_\sigma : c_1$	-0.050 ± 0.021	1.360 ± 0.028	-0.014 ± 0.015	0.929 ± 0.020
μ	-0.026 ± 0.023	1.517 ± 0.033	-0.011 ± 0.016	1.017 ± 0.020

Table C.1: Mean and sigma of a Gaussian fit to the pull distributions of $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, f_2, σ_σ: c_0, σ_σ: c_1, and μ. The columns labelled with "cor." are after correction of the covariance matrix as described in the text.

The use of sWeights in the likelihood fit of the decay-time distribution of prompt background candidates results in a wrong uncertainties on the estimated parameters [71, p. 172, Ch. 8], as noted in section 4.2. To study the effect in the context of the resolution model calibration, a series of pseudo-experiments is performed, where the $\mu^+\mu^-$ invariant mass distribution and decay-time distributions of background containing a $J/\psi \to \mu^+\mu^-$ decay and combinatorial background are generated. sWeights are then calculated as described above, followed by a weighted likelihood fit of the decay-time distribution.

The parameters estimated by the fit are stored and using the generated values, the pull is calculated, which is given by:

$$\mathrm{p} = \frac{\nu - \nu_{\mathrm{true}}}{\sigma_\nu}, \tag{C.2}$$

where p is the pull, ν is the estimated value a parameter, σ_ν is it's estimated uncertainty and ν_{true} its true value. If ν and σ_ν are estimated correctly, the distribution of p is a Gaussian with mean 0 and width 1.

The pull distribution for the resolution model parameters $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, f_2, σ_σ: c_0, σ_σ: c_1, and μ are shown in fig. C.1. The means and widths of a Gaussian PDF fit to these distributions are listed in table C.1. From the distributions in fig. C.1 and the first two columns in table C.1 it can be seen that there are no significant biases on the parameters, while the uncertainties are underestimated by 30–50%.

In a counting experiment with weighted events, the variance of the estimated parameters is given by $\sum W_i^2$, where W_i are the event weights, while a maximum likelihood fit will estimate the uncertainty to be $\sum W_i$ [81]. To correct for this difference, the covariance matrix is divided by $\alpha = \sum W_i / \sum W_i^2$. To test this procedure, an additional set of pseudo-experiments including the correction is

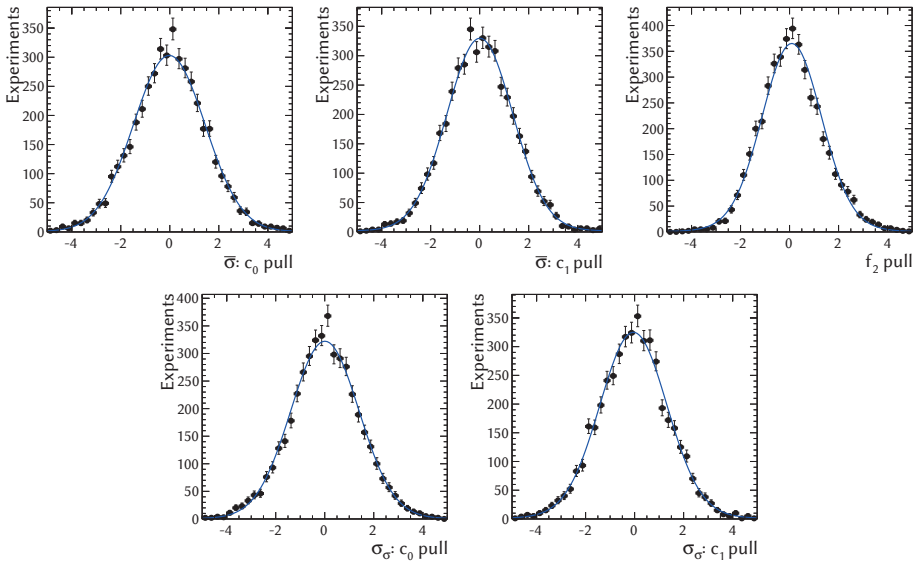


Figure C.1: Pulls of $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, f_2, σ_σ: c_0, σ_σ: c_1, and μ obtained from 4500 pseudo-experiments.

performed. The pull distributions of the resulting parameter estimates and the result of a fit to the pull distribution of a Gaussian PDF are shown in fig. C.1 and given in the two right-most columns of table C.1. From these it can be seen that the biases are still compatible with zero.

They widths of the pull distributions are not compatible with unity in all cases, which indicates that while the correction improves the estimated uncertainties, it is not perfect. As described in section 6.3.4, the uncertainties on the resolution model parameters are sufficiently small to give a negligible contribution to both the statistical and systematic uncertainties on the physics parameters of interest. The exercise presented in this appendix is nevertheless interesting, since it has clarified the amount by which the uncertainties are underestimated and that applying the correction α is beneficial.

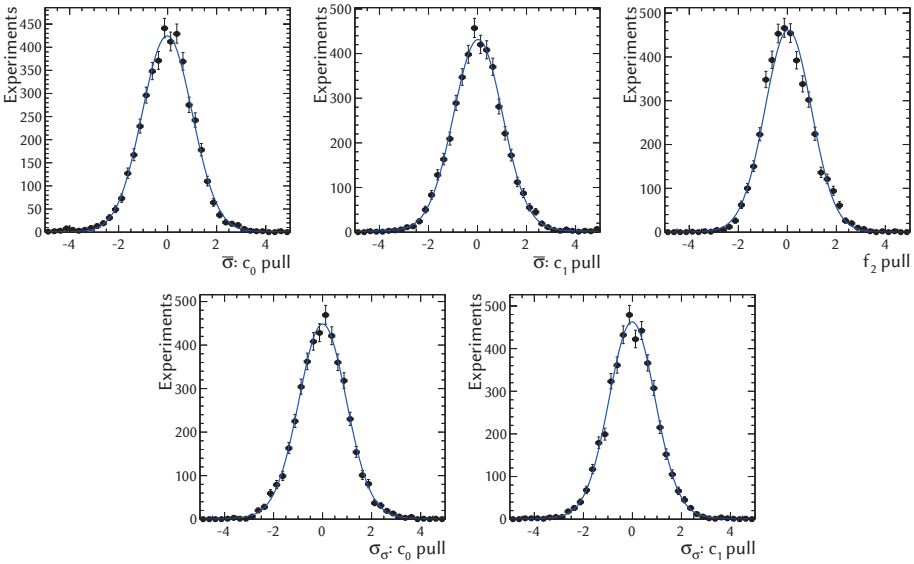


Figure C.2: Pulls of $\overline{\sigma}$: c_0, $\overline{\sigma}$: c_1, f_2, σ_σ: c_0, σ_σ: c_1, and μ obtained from 3000 pseudo-experiments after correction of the covariance matrix.

Appendix D

Reparameterisation of a Double Gaussian

The decay–time resolution model used in the analysis of ϕ_s has been defined in eq. (4.13); a simplified version without per–event decay–time errors is used here for clarity:

$$\mathcal{R}(t) = \sum_{i=1}^{n} \frac{f_i}{\sqrt{2\pi}\sigma_i} \mathrm{e}^{\frac{(t-\mu)^2}{2\sigma_i^2}}. \tag{D.1}$$

The parameterisation of the model as given by eq. (D.1) yields large correlations between the widths of the Gaussians and their relative fractions after a successful maximum–likelihood fit to the data. As described in section 6.2.1, a double Gaussian resolution model has been chosen and for this model an alternative parameterisation has been designed, which reduces the correlations between the parameters mentioned. While this reparameterisation has been done in the context of a resolution model, it may be of interest to any double–Gaussian PDF — with a common mean for all components — being used to fit a data sample.

The ansatz for the alternative parameterisation comes from the first two moments of a double Gaussian shape with a common mean: the average and the RMS of the two widths, while leaving the fraction of the second Gaussian unchanged. The squared RMS is used to avoid roots in the inverse parameterisation, which might be problematic for the minimisation algorithm by creating a

set of valid but unphysical parameter values.

$$\overline{\sigma} = (1-f)\sigma_1 + f\sigma_2 \tag{D.2}$$

$$\mathrm{RMS}_\sigma^2 = (1-f)\sigma_1^2 + f\sigma_2^2 \tag{D.3}$$

To obtain the original parameters in terms of the new parameters, the above equations are inverted by expressing i.e. σ_1 in terms of $\overline{\sigma}$ and σ_2 using eq. (D.2), substituting this into eq. (D.3) and solving the result for σ_2:

$$\sigma_1 = \frac{\overline{\sigma} - f\sigma_2}{1-f}, \tag{D.4}$$

$$\mathrm{RMS}_\sigma^2 = (1-f)\left(\frac{\overline{\sigma} - f\sigma_2}{1-f}\right)^2 + f\sigma_2^2, \tag{D.5}$$

$$= \frac{1}{1-f}\left(\overline{\sigma}^2 - 2f\overline{\sigma}\sigma_2 + f\sigma_2^2\right), \tag{D.6}$$

$$= \frac{1}{1-f}\left(f(\sigma_2 - \overline{\sigma})^2 + (1-f)\overline{\sigma}^2\right), \tag{D.7}$$

$$\sigma_2 = \pm\sqrt{\frac{1-f}{f}(\mathrm{RMS}_\sigma^2 - \overline{\sigma}^2)} + \overline{\sigma}. \tag{D.8}$$

Similar steps using σ_1 instead of σ_2 yield:

$$\sigma_1 = \pm\sqrt{\frac{f}{1-f}(\mathrm{RMS}_\sigma^2 - \overline{\sigma}^2)} + \overline{\sigma}. \tag{D.9}$$

Both the equations for σ_1 and σ_2 contain the term $\mathrm{RMS}_\sigma^2 - \overline{\sigma}^2$, which is recognised as the standard error of the "distribution" of these two widths. Replacing $\mathrm{RMS}_\sigma^2 - \overline{\sigma}^2$ by σ yields the final parameterisation; the sign before the square–root term in eq. (D.9) is chosen such that the first Gaussian has a width smaller than the average, while the second Gaussian has a width larger than the average.

$$\sigma_1 = -\sqrt{\frac{f}{1-f}}\sigma + \overline{\sigma}, \tag{D.10}$$

$$\sigma_2 = \sqrt{\frac{1-f}{f}}\sigma + \overline{\sigma} \tag{D.11}$$

and its inverse form is given by:

$$\overline{\sigma} = (1 - f)\sigma_1 + f\sigma_2, \tag{D.12}$$

$$\sigma = \sqrt{f(1 - f)}\,(\sigma_1 - \sigma_2)\,. \tag{D.13}$$

Bibliography

[1] A. Einstein. *Zur Elektrodynamik bewegter Körper*. In: *Annalen der Physik* 322.10 (1905), pp. 891–921. ISSN: 1521-3889. DOI: 10.1002/andp.19053221004 (cit. on p. 2).

[2] P. A. M. Dirac. *The Quantum Theory of the Electron*. In: *Royal Society of London Proceedings Series A* 117 (Feb. 1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023 (cit. on p. 2).

[3] C. D. Anderson. *The Positive Electron*. In: *Phys. Rev.* 43 (6 Mar. 1933), pp. 491–494. DOI: 10.1103/PhysRev.43.491 (cit. on p. 2).

[4] G. Lemaître. *Expansion of the universe, A homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae*. In: MNRAS 91 (Mar. 1931), pp. 483–490 (cit. on p. 2).

[5] E. Hubble. *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*. In: *Proceedings of the National Academy of Science* 15 (Mar. 1929), pp. 168–173. DOI: 10.1073/pnas.15.3.168 (cit. on p. 2).

[6] AMS Collaboration, M. Aguilar et al. *The Alpha Magnetic Spectrometer (AMS) on the International Space Station: Part I - results from the test flight on the space shuttle*. In: Phys. Rep. 366 (Aug. 2002), pp. 331–405. DOI: 10.1016/S0370-1573(02)00013-3 (cit. on p. 2).

[7] A.D. Sakharov. *Violation of CP invariance, C asymmetry, and baryon asymmetry in the universe*. In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35 (cit. on p. 3).

[8] ATLAS Collaboration, Georges Aad et al. *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*. In: *Phys.Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex] (cit. on p. 3).

[9] CMS Collaboration, Serguei Chatrchyan et al. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. In: *Phys.Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex] (cit. on p. 3).

[10] P. Higgs. *Broken Symmetries and the Masses of Gauge Bosons*. In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508 (cit. on p. 3).

[11] F. Englert and R. Brout. *Broken Symmetry and the Mass of Gauge Vector Mesons*. In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321 (cit. on p. 3).

[12] C. S. Wu et al. *Experimental Test of Parity Conservation in Beta Decay*. In: *Physical Review* 105 (Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413 (cit. on p. 3).

[13] CKMfitter Group, J. Charles et al. *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*. In: *Eur. Phys. J.* C41 (2005), pp. 1–131. DOI: 10.1140/epjc/s2005-02169-1. arXiv: hep-ph/0406184 [hep-ph]. Updated results and plots available at: http://ckmfitter.in2p3.fr (cit. on pp. 3, 16, 17, 169, 170).

[14] Particle Data Group, K.A. Olive. et al. *Review of Particle Physics*. In: *Chin. Phys. C* 38.9 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001 (cit. on pp. 6, 16, 20, 70, 75, 79, 86, 101).

[15] H.-Y. Cheng. *The Strong CP Problem Revisited*. In: *Phys. Rept.* 158 (1988), p. 1. DOI: 10.1016/0370-1573(88)90135-4 (cit. on p. 6).

[16] C.A. Baker et al. *An Improved experimental limit on the electric dipole moment of the neutron*. In: *Phys. Rev. Lett.* 97 (2006), p. 131801. DOI: 10.1103/PhysRevLett.97.131801 (cit. on p. 6).

[17] N. Cabibbo. *Unitary Symmetry and Leptonic Decays*. In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531 (cit. on p. 6).

[18] M. Kobayashi and T. Maskawa. *CP Violation in the Renormalizable Theory of Weak Interaction*. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652 (cit. on p. 6).

[19] J. H. Christenson et al. *Evidence for the 2π Decay of the K_2^0 Meson*. In: *Phys. Rev. Lett.* 13 (July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138 (cit. on p. 8).

[20] LHCb Collaboration, R. Aaij. *First evidence of direct CP violation in charmless two-body decays of B_s mesons*. In: *Phys. Rev. Lett.* 108 (Feb. 2012), 201601. 8 p (cit. on p. 8).

[21] V. Weisskopf and E. Wigner. *Über die natürliche Linienbreite in der Strahlung des harmonischen Oszillators*. In: *Z. Phys.* 65 (1930), pp. 18–29. DOI: 10.1007/BF01397406 (cit. on p. 9).

[22] V. Weisskopf and Eugene P. Wigner. *Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie*. In: *Z. Phys.* 63 (1930), pp. 54–73. DOI: 10.1007/BF01336768 (cit. on p. 9).

[23] P.K. Kabir. *The CP Puzzle: Strange Decays of the Neutral Kaon*. Academic Press, 1968 (cit. on p. 9).

[24] D. van Eijk. *Ageing and the Decay of Beauty*. PhD thesis. V. Universiteit, Amsterdam, 2012 (cit. on pp. 9, 12, 15, 95, 97).

[25] T. du Pree. *Search for a Strange Phase in Beautiful Oscillations*. PhD thesis. Amsterdam: Vrije U., Amsterdam, 2010 (cit. on pp. 9, 12, 97).

[26] LHCb collaboration, R. Aaij et al. *Measurement of the flavour-specific CP-violating asymmetry a^s_{sl} in B_s decays*. In: *Phys. Lett. B* 728 (Aug. 2013), 607–615. 25 p (cit. on p. 9).

[27] LHCb collaboration, R. Aaij et al. *Measurement of the semileptonic CP asymmetry in B^0 –$\overline{B}^0$ mixing*. In: (2014). arXiv: 1409.8586 [hep-ex] (cit. on p. 9).

[28] S. Ali. personal communication. Sept. 1, 2014 (cit. on p. 13).

[29] LHCb collaboration, R. Aaij et al. *Differential branching fraction and angular analysis of the decay $B_s \to \varphi\mu^+\mu^-$*. In: *JHEP* 1307 (2013), p. 084. DOI: 10.1007/JHEP07(2013)084. arXiv: 1305.2168 [hep-ex] (cit. on p. 15).

[30] L. Wolfenstein. *Parametrization of the Kobayashi-Maskawa Matrix*. In: *Phys. Rev. Lett.* 51 (21 Nov. 1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945 (cit. on pp. 15, 17).

[31] M. A. Shifman. *Foreword to ITEP lectures in particle physics*. In: (1995). arXiv: hep-ph/9510397 [hep-ph] (cit. on p. 16).

[32] S. Faller, R. Fleischer, and T. Mannel. *Precision physics with $B_s \to J/\psi\varphi$ at the LHC: The quest for new physics*. In: *Phys. Rev. D* 79 (1 Jan. 2009), p. 014005. DOI: 10.1103/PhysRevD.79.014005 (cit. on p. 17).

[33] R. Aaij. personal communication. Oct. 1, 2014 (cit. on p. 19).

[34] R. Aaij et al. *Performance of the LHCb Vertex Locator*. In: *JINST* 9 (2014), p. 09007. DOI: 10.1088/1748-0221/9/09/P09007. arXiv: 1405.7808 [physics.ins-det] (cit. on p. 23).

[35] LHCb Collaboration, P. R. Barbosa-Marinho et al. *LHCb VELO (VErtex LOcator): Technical Design Report*. Geneva: CERN, 2001 (cit. on p. 24).

[36] LHCb collaboration, R. Aaij et al. *LHCb Detector Performance*. In: *Int. J. Mod. Phys. A* (2014) (cit. on pp. 26, 27, 31, 34).

[37] R. E. Kalman. *A New Approach to Linear Filtering and Prediction Problems*. In: *Transactions of the ASME–Journal of Basic Engineering* 82.Series D (1960), pp. 35–45 (cit. on pp. 28, 60).

[38] C. Fitzpatrick. *A measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*. Presented 24 May 2012. PhD thesis. Edinburgh U., Apr. 2012 (cit. on p. 29).

[39] P.A. Cherenkov. *Visible Emission of Clean Liquids by Action of γ Radiation*. In: *Doklady Akad. Nauk SSSR* 2 (1934), p. 451 (cit. on p. 29).

[40] LHCb Collaboration, R. Aaij et al. *Measurement of the ratio of branching fractions BR($B^0 \to K^{*0}\gamma$)/BR($B_s^0 \to \varphi\gamma$) and the direct CP asymmetry in $B^0 \to K^{*0}\gamma$*. In: *Nucl. Phys.* B867 (2013), pp. 1–18. DOI: 10.1016/j.nuclphysb.2012.09.013. arXiv: 1209.0313 [hep-ex] (cit. on p. 32).

[41] A. Bazilevsky et al. *Performance of the PHENIX EM Calorimeter*. In: *Nuclear Science, IEEE Transactions on* 43.3 (June 1996), pp. 1491–1495. DOI: 10.1109/23.507090 (cit. on p. 33).

[42] J. Badier et al. *Shashlik calorimeter: Beam test results*. In: *Nucl. Instrum. Meth.* A348 (1994), pp. 74–86. DOI: 10.1016/0168-9002(94)90844-3 (cit. on p. 33).

[43] LHCb Collaboration, P. R. Barbosa-Marinho et al. *LHCb muon system: Technical Design Report*. Geneva: CERN, 2001 (cit. on pp. 34, 35).

[44] P. Koppenburg. *Nu, Mu and Pile-Up. The LHCb definitions of what we see and what we don't see*. 2010. URL: https://twiki.cern.ch/twiki/bin/view/LHCb/NuMuPileUp (cit. on pp. 35, 81).

[45] J. Laubser. *Conception et réalisation de l'unité de décision du système de déclenchement de premier niveau du détecteur LHCb au LHC*. Presented on 29 Nov 2007. PhD thesis. Clermont-Ferrand: U. Blaise Pascal - Clermont II, 2007 (cit. on p. 37).

[46] LHCb Collaboration, R. Antunes-Nobrega et al. *LHCb trigger system: Technical Design Report*. revised version number 1 submitted on 2003-09-24 12:12:22. Geneva: CERN, 2003 (cit. on pp. 37, 38).

[47] R. Aaij et al. *The LHCb Trigger and its Performance in 2011*. In: *JINST* 8 (2013), P04022 (cit. on p. 38).

[48] S. Oggero. *Beauty in the Crowd: Commissioning of the LHCb Pile-Up detector and First evidence of* $B_s^0 \to \mu^+\mu^-$. PhD thesis. VU Univ. Amsterdam, Apr. 2013 (cit. on p. 38).

[49] LHCb Collaboration, R. Antunes-Nobrega et al. *LHCb computing: Technical Design Report*. Submitted on 11 May 2005. Geneva: CERN, 2005 (cit. on p. 38).

[50] LHCb Collaboration. R. Graciani Diaz. *LHCb Computing Resources: 2011 re-assessment, 2012 request and 2013 forecast*. LHCb-PUB-2011-009.

[51] LHCb Collaboration. R. Graciani Diaz. *LHCb Computing Resources: 2012 re-assessment, 2013 request and 2014 forecast*. LHCb-PUB-2012-004.

[52] LHCb collaboration, R. Aaij et al. *Prompt charm production in pp collisions at* $\sqrt{s} = 7$ *TeV*. In: *Nucl. Phys.* B871 (2013), pp. 1–20. DOI: 10.1016/j.nuclphysb.2013.02.010. arXiv: 1302.2864 [hep-ex] (cit. on p. 40).

[53] LHCb Collaboration, R. Aaij et al. *Measurement of* J/ψ *production in pp collisions at* $\sqrt{s} = 7$ *TeV*. In: *Eur. Phys. J.* C71 (2011), p. 1645. DOI: 10.1140/epjc/s10052-011-1645-y. arXiv: 1103.0423 [hep-ex] (cit. on p. 40).

[54] LHCb Collaboration. O. Callot. *FastVelo, a fast and efficient pattern recognition package for the Velo*. LHCb-PUB-2011-001.

[55] B. Hommels. *The tracker in the trigger of LHCb*. PhD thesis. VU Univ. Amsterdam, 2006 (cit. on p. 50).

[56] R. O. Duda and P. E. Hart. *Use of the Hough Transformation to Detect Lines and Curves in Pictures*. In: *Commun. ACM* 15.1 (Jan. 1972), pp. 11–15. DOI: 10.1145/361237.361242 (cit. on p. 59).

[57] F. Archilli et al. *Performance of the Muon Identification at LHCb*. In: *JINST* 8 (2013), P10020. DOI: 10.1088/1748-0221/8/10/P10020. arXiv: 1306.0249 [physics.ins-det] (cit. on p. 60).

[58] J. Gaiser. *Charmonium Spectroscopy From Radiative Decays of the* J/ψ *and* ψ'. In: SLAC-R-255 (1982). p. 178 (cit. on pp. 64, 122).

[59] T. Skwarnicki. *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*. PhD thesis. Institute of Nuclear Physics, Krakow, 1986. App. E (cit. on pp. 64, 122).

[60] LHCb HLT project, J. Albrecht et al. *Performance of the LHCb High Level Trigger in 2012*. In: *J. Phys. Conf. Ser.* 513 (2014), p. 012001. DOI: 10.1088/1742-6596/513/1/012001 (cit. on pp. 67, 68, 72).

[61] B. P. Roe et al. *Boosted decision trees as an alternative to artificial neural networks for particle identification*. In: *Nucl. Instrum. Meth.* A543 (2005), pp. 577–584. DOI: 10.1016/j.nima.2004.12.018. arXiv: physics/0408124 [physics] (cit. on p. 71).

[62] V. V. Gligorov and M. Williams. *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*. In: *JINST* 8 (2013), P02013. DOI: 10.1088/1748-0221/8/02/P02013. arXiv: 1210.6861 [physics.ins-det] (cit. on p. 71).

[63] SLD Collaboration, K. Abe et al. *A Measurement of R(b) using a vertex mass tag*. In: *Phys. Rev. Lett.* 80 (1998), pp. 660–665. DOI: 10.1103/PhysRevLett.80.660 (cit. on p. 72).

[64] W. Hulsbergen. *Decay chain fitting with a Kalman filter*. In: *Nucl. Instrum. Meth.* A552 (2005), pp. 566–575. DOI: 10.1016/j.nima.2005.06.078. arXiv: physics/0503191 [physics] (cit. on p. 75).

[65] D. Martinez Santos. *Apolonius function for mass fitting*. In: *Nucl. Instrum. Meth.* (2014). In preparation (cit. on p. 76).

[66] LHCb Collaboration. M. Needham. *Clone Track Identification using the Kullback-Liebler Distance*. LHCb-2008-002.

[67] O. S. Brüning et al. *LHC Design Report*. Vol. 1. Geneva: CERN, 2004 (cit. on p. 81).

[68] LHCb collaboration, R. Aaij et al. *Measurement of B meson production cross-sections in proton-proton collisions at* $\sqrt{s} = 7$ TeV. In: *JHEP* 08 (2013), p. 117. DOI: 10.1007/JHEP08(2013)117. arXiv: 1306.3663 [hep-ex] (cit. on p. 81).

[69] LHCb Collaboration. R. Aaij et al. *Flavour tagged time-dependent angular analysis of* $B_s \to J/\psi K^+K^-$ *decays in the low* K^+K^- *mass range*. LHCb-ANA-2014-039.

[70] M. Pivk and F. R. Le Diberder. *SPlot: A Statistical tool to unfold data distributions*. In: *Nucl. Instrum. Meth.* A555 (2005), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106. arXiv: physics/0402083 [physics.data-an] (cit. on p. 84).

[71] W. T. Eadie et al. *Statistical methods in experimental physics*. Amsterdam: North-Holland Pub. Co.; New York: American Elsevier Pub. Co, 1971. ISBN: 0720402395 (North-Holland) (cit. on pp. 85, 177).

[72] G. Breit and E. Wigner. *Capture of Slow Neutrons*. In: *Phys. Rev.* 49 (7 Apr. 1936), pp. 519–531. DOI: 10.1103/PhysRev.49.519 (cit. on p. 86).

[73] LHCb Collaboration, R. Aaij et al. *Amplitude analysis and the branching fraction measurement of* $B_s \to J/\psi K^+ K^-$. In: *Phys. Rev. D* 87 (Feb. 2013), 072004. 41 p (cit. on p. 87).

[74] LHCb collaboration, A. A. Alves Jr. et al. *The LHCb detector at the LHC*. In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005 (cit. on pp. 90, 92, 107).

[75] H. G. Moser and A. Roussarie. *Mathematical Methods for* $B^0\,\overline{B}^0$ *Oscillation Analyses*. In: *Nucl. Instrum. Meth. Phys. Res., A* 384.CERN-OPEN-99-030. (May 1996), 491–505. 27 p (cit. on p. 93).

[76] W. Hulsbergen. personal communication. Mar. 20, 2013 (cit. on p. 94).

[77] S. Wandernoth. *Measurement of the* B_s-$\overline{B}_s$ *Oscillation Frequency at LHCb using* $1fb^{-1}$ *of data taken in 2011*. PhD thesis. Heidelberg University, 2014. URL: http://www.ub.uni-heidelberg.de/archiv/16316 (cit. on p. 94).

[78] M. Dorigo. personal communication. May 1, 2014 (cit. on pp. 99, 100).

[79] LHCb collaboration, Roel Aaij et al. *Measurements of the* B^+, B^0, B_s^0 *meson and* Λ_b^0 *baryon lifetimes*. In: *JHEP* 1404 (2014), p. 114. DOI: 10.1007/JHEP04(2014)114. arXiv: 1402.2554 [hep-ex] (cit. on pp. 108, 159).

[80] LHCb Collaboration, R. Aaij et al. *Precision measurement of CP violation in* $B_s \to J/\psi K^+ K^-$ *decays*. In: (2014). arXiv: 1411.3104 [hep-ex] (cit. on pp. 109, 160, 165, 170).

[81] J. van Leerdam. personal communication. May 1, 2014 (cit. on pp. 112, 158–161, 177).

[82] R. Bailey et al. *Measurement of the lifetime of charged and neutral D mesons with high resolution silicon strip detectors*. In: *Z.Phys.* C28 (1985), pp. 357–363. DOI: 10.1007/BF01413598 (cit. on p. 112).

[83] J. Rademacker. *Reduction of statistical power per event due to upper lifetime cuts in lifetime measurements.* In: *Nucl. Instrum. Meth.* A570 (2007), pp. 525–528. DOI: 10.1016/j.nima.2006.09.090. arXiv: hep-ex/0502042 [hep-ex] (cit. on p. 112).

[84] CDF Collaboration, T. Aaltonen et al. *Measurement of the B^- lifetime using a simulation free approach for trigger bias correction.* In: *Phys. Rev. D* 83 (3 Feb. 2011), p. 032008. DOI: 10.1103/PhysRevD.83.032008 (cit. on p. 112).

[85] CDF Collaboration, R. Blair et al. *The CDF-II detector: Technical design report.* Tech. rep. Fermilab, 1996 (cit. on p. 114).

[86] K. S. Cranmer. *Kernel estimation in high-energy physics.* In: *Comput. Phys. Commun.* 136 (2001), pp. 198–207. DOI: 10.1016/S0010-4655(00)00243-5. arXiv: hep-ex/0011057 [hep-ex] (cit. on p. 147).

[87] LHCb Collaboration. R. Aaij et al. *Letter of Intent for the LHCb Upgrade.* CERN-LHCC-2011-001.

[88] LHCb Collaboration, R. Aaij et al. *Precision measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency with the decay $B_s \to D_s^- \pi^+$.* In: *New J. Phys.* 15 (Apr. 2013), 053021. 18 p (cit. on p. 158).

[89] LHCb collaboration, R. Aaij et al. *Observation of the Decay $B_c^+ \to B_s \pi^+$.* In: *Phys. Rev. Lett.* 111.18 (2013), p. 181801. DOI: 10.1103/PhysRevLett.111.181801 (cit. on p. 165).

[90] A. Lenz and U. Nierste. *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons.* In: (2011). arXiv: 1102.4274 [hep-ph] (cit. on pp. 169, 170).

[91] Heavy Flavor Averaging Group, Y. Amhis et al. *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014.* In: (2014). arXiv: 1412.7515 [hep-ex]. updated results and plots available at http://www.slac.stanford.edu/xorg/hfag (cit. on p. 169).

[92] CDF collaboration, T. Aaltonen et al. *Measurement of the bottom-strange meson mixing phase in the full CDF data set.* In: *Phys. Rev. Lett.* 109 (2012), p. 171802. DOI: 10.1103/PhysRevLett.109.171802. arXiv: 1208.2967 [hep-ex] (cit. on p. 170).

[93] D0 collaboration, V. M. Abazov et al. *Measurement of the CP-violating phase $\phi_s^{J/\psi\varphi}$ using the flavor-tagged decay $B_s \to J/\psi\varphi$ in 8 fb^{-1} of $p\bar{p}$ collisions.* In: *Phys. Rev.* D85 (2012), p. 032006. DOI: 10.1103/PhysRevD.85.032006. arXiv: 1109.3166 [hep-ex] (cit. on p. 170).

[94] ATLAS Collaboration, G. Aad et al. *Time-dependent angular analysis of the decay $B_s \to J/\psi\varphi$ and extraction of $\Delta\Gamma_s$ and the CP-violating weak phase ϕ_s by ATLAS*. In: *JHEP* 1212 (2012), p. 072. DOI: 10.1007/JHEP12(2012)072. arXiv: 1208.0572 [hep-ex] (cit. on p. 170).

[95] CMS Collaboration, *Measurement of the CP-violating weak phase phi-s and the decay width difference DeltaGamma-s using the $B_s \to J/\psi\varphi$ decay channel*. Tech. rep. CMS-PAS-BPH-13-012. Geneva: CERN, 2014 (cit. on p. 170).

[96] V. Belyaev. *LoKi: Smart and Friendly C++ Analysis Toolkit*. 2005–2014. URL: http://lhcb-comp.web.cern.ch/lhcb-comp/Analysis/Loki (cit. on p. 172).

Summary

If we zoom in close enough to things around us, or to ourselves, smaller and smaller structures can be identified. Our limbs and organs consist of cells, which consist of molecules which consist of atoms. Atoms, in turn, consist of a nucleus that is surrounded by a cloud of electrons, and the nucleus contains electrically charged protons and neutral neutrons. For a long time, protons and neutrons — like electrons — were thought to be elemental particles, particles that are not made up of other particles. At the end of the sixties, however, it became clear that protons and neutrons do contain other particles, called quarks [1].

In 1928, Paul Dirac proposed that for every particle, a so called anti-particle should exist [2], that has identical mass but opposite charge. In 1932 the first anti-particle was discovered, the anti-particle of the electron: the positron [3]. It has subsequently become clear that for each particle, a corresponding anti-particles exists.

In addition to particles, four fundamental interactions between particles exist in nature: gravity, the electromagnetic interaction and the weak and strong nuclear forces. An important property of the electromagnetic and strong nuclear interactions is that they act in the same way on particles and anti-particles (C symmetry) and also on particles and particles whose spacial coordinates have been inverted ($\mathbf{x} \rightarrow$ -$\mathbf{x}$), i.e. for particles and their mirror image (P symmetry).

Until 1956, this was assumed to also hold for the weak nuclear force, although this had not been verified experimentally. When a measurement was proposed by Lee and Yang [4] and performed the same year [5], it was discovered that the weak force treats particles and their mirror image maximally differently. The combination of the C and P symmetries, CP symmetry was then proposed to be a conserved quantity, i.e. all forces are the same for particles and mirrored anti-particles. In 1964 it was shown that this was also not true, when the probability that a neutral kaon transforms into a neutral anti-kaon was observed to be different from the probability of the reversed process [6].

Originating from the sixties, a model has been created that describes the electromagnetic interaction and the weak and strong nuclear forces: the so-called Standard Model of particle physics. In the Standard Model, the possibility for *CP* violation to occur originates from the presence of a single complex-valued parameter, η. The value of this parameter influences many observable processes, which implies that as soon as the value of η has been measured with a first set of processes, predictions can be made for other processes. Using such a process, a measurement that is incompatible with its prediction with sufficient significance, is direct proof of the existence of physics beyond the Standard Model. The measurement described in this thesis is such a measurement.

This measurement is performed by measuring a single observable, ϕ_s, which is related to η, and whose value has been precisely predicted by the Standard Model to be close to zero [7]. To determine the value of ϕ_s, the decay-time and decay-angle dependent decay probability of B_s^0 mesons to a J/ψ and a meson is measured. A meson is a particle that consists of two quarks. The decay-time of a B_s^0 meson is the time between its production and decay as measured in its rest frame. The decay angles are defined by the directions of the final decay products: two charged muons ($\mu^+\mu^-$) and two charged kaons (K^+K^-).

B_s^0 mesons are produced in collisions between protons in the Large Hadron Collider at Cern. Since these mesons are predominantly produced in directions close to the direction of the proton beams, the LHCb detector has been optimised to measure the tracks and properties of particles, such as muons and kaons originating from a $B_s \to J/\psi\,\varphi$ decay, in the forward direction.

Throughout 2011 and 2012, the proton beams crossed each other approximately 15 million times per second. Of those 15 million crossings per second, approximately 12 million resulted in at least one collision between two protons. The detector produces about 60 kiB of data for each such event, which results in a total data rate of 700 GiB per second. Only a small fraction of all proton-proton collisions results in the production of a B_s^0 or $\overline{B}_s^0$ meson and only a fraction of those produced mesons decays according to $B_s \to J/\psi\,\varphi$. It is therefore not necessary to store all data, and to store only data from interesting crossings, it is filtered in two consecutive steps. The first of these filters is implemented in programmable chips and the second is a software application running on a farm of PC servers. This second filtering step is called the High Level Trigger.

The HLT is again divided in two stages, where the first stage (HLT1) is based on general selection criteria, such as the presence of two reconstructed muons with the same origin, to reduce the number of crossings that must be analysed to about eighty thousand per second with a minimal expense of CPU time. The

second stage (HLT2) is then allowed to use more CPU time to obtain additional information that can be used to select data from interesting crossings with the highest possible efficiency.

Data that has been selected by the HLT is further processed before the statistical analysis of the decay-rate can be performed to determine ϕ_s. First, all tracks of charged particles are reconstructed and from those, suitable muon and kaons tracks are selected, that are combined to form candidate decays. A random combination of pairs of oppositely charged muons and kaons can resemble a real $B_s \to J/\psi\,\varphi$ decay to such an extent that it is accepted by the selection. Such candidates are not suitable for analysis, since they do not correspond to a real $B_s \to J/\psi\,\varphi$ decay, and they are removed using a statistical subtraction method.

To measure ϕ_s, maximum likelihood estimation is employed. Given the data, this method determines according to the theoretical decay-time and decay-angle dependent probability density of the $B_s \to J/\psi\,\varphi$ process, the most likely value of ϕ_s and its uncertainty.

Before the most likely value of ϕ_s can be determined, the theoretical model must be augmented to take the effects of experimental uncertainties on the measured quantities (decay-time and angles) and selection effects into account. The fact that the probability to reconstruct and select a $B_s \to J/\psi\,\varphi$ decay depends on its decay time, must, for example, be taken into account. In addition, the decay-time uncertainty model must be calibration and confirmed to be correct using real and simulated data.

Once these, and other, effects have been accounted for, the most likely estimate of ϕ_s and its uncertainty can be determined; it is given by:

$$\phi_s = \text{-}0.057 \pm 0.051 \pm 0.007 \quad \text{rad},$$

where the statistical and systematic uncertainties are given separately. The statistical uncertainty of this measurements is larger than the systematic uncertainty, which implies that the larger quantities of data that will be recorded by the LHCb detector during Run II of the LHC and by the upgraded detector after 2017, can be used to further reduce the uncertainty on the measurement to match the uncertainty on its Standard Model prediction.

The measurement of ϕ_s is compatible with its Standard Model prediction, which is given by [7]:

$$\phi_s = \text{-}0.0363^{+0.0014}_{-0.0012} \quad \text{rad}, \tag{D.14}$$

and thus there is, currently, no evidence for physics beyond the Standard Model.

References

[1] E. D. Bloom et al. *High-Energy Inelastic e–p Scattering at 6° and 10°*. In: *Phys. Rev. Lett.* 23 (16 Oct. 1969), pp. 930–934. DOI: 10.1103/PhysRevLett.23.930. URL: http://link.aps.org/doi/10.1103/PhysRevLett.23.930 (cit. on p. 194).

[2] P. A. M. Dirac. *The Quantum Theory of the Electron*. In: *Royal Society of London Proceedings Series A* 117 (Feb. 1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023 (cit. on p. 194).

[3] C. D. Anderson. *The Positive Electron*. In: *Phys. Rev.* 43 (6 Mar. 1933), pp. 491–494. DOI: 10.1103/PhysRev.43.491 (cit. on p. 194).

[4] T. D. Lee and C. N. Yang. *Question of Parity Conservation in Weak Interactions*. In: *Phys. Rev.* 104 (1 Oct. 1956), pp. 254–258. DOI: 10.1103/PhysRev.104.254. URL: http://link.aps.org/doi/10.1103/PhysRev.104.254 (cit. on p. 194).

[5] C. S. Wu et al. *Experimental Test of Parity Conservation in Beta Decay*. In: *Physical Review* 105 (Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413 (cit. on p. 194).

[6] J. H. Christenson et al. *Evidence for the 2π Decay of the K_2^0 Meson*. In: *Phys. Rev. Lett.* 13 (July 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138 (cit. on p. 194).

[7] CKMfitter Group, J. Charles et al. *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*. In: *Eur. Phys. J.* C41 (2005), pp. 1–131. DOI: 10.1140/epjc/s2005-02169-1. arXiv: hep-ph/0406184 [hep-ph]. Updated results and plots available at: http://ckmfitter.in2p3.fr (cit. on pp. 195, 196).

Samenvatting

Geactiveerd door *CP*-schending

Realtime Selectie en Reconstructie van Vervallen van het Type $B_s \rightarrow J/\psi\,\varphi$

Als we ver genoeg inzoomen op dingen om ons heen of op ons zelf, zien we steeds kleinere structuren. Onze ledematen en organen bestaan uit cellen, die cellen bestaan uit moleculen en die moleculen bestaan weer uit atomen. Atomen bestaan op hun beurt uit een kern, met daaromheen een wolk elektronen, en de kern bevat elektrisch geladen protonen en neutrale neutronen. Lang werd gedacht dat protonen en neutronen — net als elektronen — elementaire deeltjes waren, deeltjes die niet uit andere deeltjes bestaan. Eind jaren zestig werd echter duidelijk dat protonen en neutronen wel degelijk uit andere deeltjes bestaan, de zogenaamde quarks [1].

In 1928 stelde Paul Dirac voor dat er voor elk deeltje een zogenaamd antideeltje bestaat [2], dat dezelfde massa maar een tegenovergestelde lading heeft. In 1932 werd als eerste antideeltje, het antideeltje van het elektron ontdekt: het positron [3]. Daarna is duidelijk geworden dat er voor alle deeltjes een antideeltje bestaat.

Naast deeltjes bestaan er in de natuur vier fundamentele wisselwerkingen tussen elementaire deeltjes: zwaartekracht, elektromagnetische kracht, zwakke kernkracht en sterke kernkracht. Een belangrijke eigenschap van de elektromagnetische en sterke kernkrachten is dat ze op deeltjes en op antideeltjes precies hetzelfde effect hebben (*C* symmetrie). Dit geldt ook voor deeltjes en deeltjes waarvan alle ruimtelijke coördinaten zijn geïnverteerd ($\mathbf{x} \rightarrow$ -$\mathbf{x}$), met andere woorden voor deeltjes en hun spiegelbeeld (*P* symmetrie).

Voor de zwakke kernkracht werd tot 1956 aangenomen dat dit ook het gevel was, hoewel dat tot dan toe niet experimenteel geverifieerd was. Toen de daartoe door Lee en Yang in dat jaar voorgestelde meting [4] werd uitgevoerd

[5], bleek dat de zwakke kernkracht niet hetzelfde is voor deeltjes en hun spiegelbeeld, maar zelfs maximaal verschillend. Vervolgens werd voorgesteld dat de combinatie van de *C* en *P* symmetrie, de *CP* symmetrie, wel behouden zou zijn; met andere woorden dat alle krachten zich hetzelfde gedragen voor deeltjes en gespiegelde antideeltjes. In 1964 bleek ook dit niet het geval te zijn toen werd gemeten dat de kans dat een neutraal kaon in een anti-kaon overgaat niet precies gelijk is aan de kans dat het omgekeerde proces plaatsvindt [6].

Vanaf de jaren zestig is er een model opgesteld dat de elektromagnetische kracht, zwakke kernkracht en sterke kernkracht beschrijft: het Standaard Model van de deeltjes fysica. *CP* symmetrie schending vindt in dit model zijn oorsprong in de aanwezigheid van één complexe parameter, η. De waarde van deze parameter beïnvloedt vele waarneembare processen, en dit betekent dat zodra de waarde van η is gemeten met behulp van bepaalde processen, er voorspellingen kunnen worden gedaan voor andere processen. Als een meting aan een dergelijk proces met voldoende zekerheid niet in overeenstemming is met haar voorspelling, is dat direct bewijs voor de onvolledigheid van het Standaard Model en het bestaan van fysica die er niet door beschreven wordt. De meting die in dit proefschrift beschreven wordt, is een dergelijke meting.

Deze meting wordt verricht door het bepalen van een parameter, ϕ_s, die gerelateerd is aan η en waarvan het Standaard Model zeer nauwkeurig voorspelt dat de waarde dicht bij nul ligt [7]. Om ϕ_s te bepalen word de vervaltijd- en hoekafhankelijke vervalkans van B_s^0 mesonen naar een J/ψ en een φ meson gemeten. Een meson is een deeltje dat bestaat uit twee quarks. De vervaltijd die wordt gebruikt, is de tijd tussen productie en verval van het B_s^0 meson, bepaald in zijn ruststelsel. De gemeten hoeken zijn gedefinieerd aan de hand van de richtingen van de uiteindelijke vervalproducten van de J/ψ en mesonen: twee muonen ($\mu^+\mu^-$) en twee geladen kaonen (K^+K^-).

B_s^0 mesonen worden geproduceerd bij botsingen tussen protonen in de Large Hadron Collider op Cern. Omdat deze mesonen vooral worden geproduceerd in voor- en achterwaartse richtingen langs de protonbundels, is de LHCb detector geoptimaliseerd om in de voorwaartse richting de sporen en eigenschappen van deeltjes, zoals de muonen en kaonen van een $B_s \to J/\psi\,\varphi$ verval, te meten.

In 2011 en 2012 kruisten de bundels elkaar ongeveer 15 miljoen keer per seconde. Van die 15 miljoen bundelkruisingen per seconde waren er ongeveer 12 miljoen waarbij ten minste één botsing tussen twee protonen plaatsvond. De detector produceerde voor elk van die kruisingen ongeveer 60 kiB aan data, wat neerkwam op ongeveer 700 GiB data per seconde. Slechts een fractie van alle proton-proton botsingen resulteert echter in de productie van een B_s^0 of $\overline{B}_s^0$ meson

en weer een fractie van de geproduceerde mesonen vervalt volgens $B_s \to J/\psi\,\varphi$, waardoor het niet nodig is om alle data op te slaan. De data van alle kruisingen wordt daarom in twee opeenvolgende stappen gefilterd, om alleen data van interessante kruisingen op te slaan. De eerste van deze twee filters is geïmplementeerd met behulp van programmeerbare chips en de tweede stap als een software applicatie die draait op PC servers. Deze tweede stap heet de High Level Trigger (Hoog Niveau Filter), of HLT.

De HLT is op zijn beurt ook verdeeld in twee stappen. De eerste stap (HLT1) is gebaseerd op algemene selectiecriteria, zoals de aanwezigheid van twee gereconstrueerde muonen met dezelfde oorsprong, om de hoeveelheid te analyseren kruisingen snel te reduceren tot ongeveer tachtigduizend per seconde. De tweede stap (HLT2) kan daarna meer tijd in beslag nemen en meer informatie gebruiken om met een zo hoog mogelijke efficiëntie data van interessante kruisingen te selecteren.

Data die door de HLT is geselecteerd wordt daarna verder verwerkt voordat de statistische analyse van de vervalkans — waar ϕ_s mee wordt bepaald — kan worden uitgevoerd. Als eerste worden alle sporen van geladen deeltjes gereconstrueerd en worden geschikte muonen en kaonen geselecteerd. Die worden vervolgens gecombineerd tot kandidaat-vervallen. Het komt voor dat een willekeurige combinatie van twee muonen en twee kaonen zodanig lijkt op een $B_s \to J/\psi\,\varphi$ verval dat de combinatie door de selectie wordt geaccepteerd. Voor de analyse zijn dit soort kandidaten ongewenst, omdat ze niet met een werkelijk $B_s \to J/\psi\,\varphi$ verval overeenkomen. Dergelijke vervallen worden met behulp van een statistische methode van de geselecteerde data afgetrokken, waardoor ze vanaf dat moment in de verdere analyse buiten beschouwing kunnen worden gelaten.

Om ϕ_s te bepalen, wordt de methode van de meest aannemelijke schatting gebruikt. Gegeven de data bepaalt deze methode de meest aannemelijke waarde van ϕ_s en haar onzekerheid aan de hand van de uit de theorie voortvloeiende kansdichtheidsmodel voor de tijd- en hoekafhankelijke vervalkans van het $B_s \to J/\psi\,\varphi$ proces.

Voordat de meest aannemelijke schatting van ϕ_s kan worden bepaald, moet het puur theoretische model worden aangepast om, onder andere, selectie-effecten en de gevolgen van meetfouten van de vervaltijd en -hoeken correct mee te nemen. Zo moet bijvoorbeeld aan het model worden toegevoegd dat de kans dat een $B_s \to J/\psi\,\varphi$ verval wordt gereconstrueerd en geselecteerd, afhangt van de vervaltijd. Daarnaast moeten de calibratie en de correctheid van het model voor de onzekerheid van de gemeten vervaltijd in detail uitgewerkt en gecontroleerd

worden aan de hand van echte en gesimuleerde data.

Zodra er rekening gehouden is met deze en andere factoren, kan de meest aannemelijke schatting van ϕ_s worden bepaald. Hij wordt gegeven door:

$$\phi_s = \text{-}0.057 \pm 0.051 \pm 0.007 \quad \text{rad},$$

waarbij de statische respectievelijk systematische onzekerheden afzonderlijk zijn gegeven. De statistische onzekerheid van deze meting is groter dan de systematische onzekerheid. Dit betekent dat de grotere hoeveelheid data die tijdens Run II van de LHC en met behulp van de verbeterde LHCb detector na 2017 zullen worden verzameld, gebruikt kunnen worden om de onzekerheid op de meting verder te verkleinen en de nauwkeurigheid van de voorspelling te evenaren.

Het resultaat van de meting van ϕ_s is in overeenstemming met de door het Standaard model voorspelde waarde van ϕ_s [7]:

$$\phi_s = \text{-}0.0363^{+0.0014}_{-0.0012} \quad \text{rad}, \tag{D.15}$$

Aan de hand van het behaalde resultaat kan het bestaan van niet door het Standaard Model beschreven fysica niet worden aangetoond.

Referenties

[1] E. D. Bloom e.a. *High-Energy Inelastic e–p Scattering at 6° and 10°*. In: *Phys. Rev. Lett.* 23 (16 okt 1969), p. 930–934. DOI: 10.1103/PhysRevLett.23.930. URL: http://link.aps.org/doi/10.1103/PhysRevLett.23.930 (zie pag. 198).

[2] P. A. M. Dirac. *The Quantum Theory of the Electron*. In: *Royal Society of London Proceedings Series A* 117 (feb 1928), p. 610–624. DOI: 10.1098/rspa.1928.0023 (zie pag. 198).

[3] C. D. Anderson. *The Positive Electron*. In: *Phys. Rev.* 43 (6 mrt 1933), p. 491–494. DOI: 10.1103/PhysRev.43.491 (zie pag. 198).

[4] T. D. Lee en C. N. Yang. *Question of Parity Conservation in Weak Interactions*. In: *Phys. Rev.* 104 (1 okt 1956), p. 254–258. DOI: 10.1103/PhysRev.104.254. URL: http://link.aps.org/doi/10.1103/PhysRev.104.254 (zie pag. 198).

[5] C. S. Wu e.a. *Experimental Test of Parity Conservation in Beta Decay*. In: *Physical Review* 105 (feb 1957), p. 1413–1415. DOI: 10.1103/PhysRev.105.1413 (zie pag. 199).

[6] J. H. Christenson e.a. *Evidence for the 2π Decay of the K_2^0 Meson.* In: *Phys. Rev. Lett.* 13 (jul 1964), p. 138–140. DOI: 10.1103/PhysRevLett.13.138 (zie pag. 199).

[7] CKMfitter Group, J. Charles e.a. *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories.* In: *Eur. Phys. J.* C41 (2005), p. 1–131. DOI: 10.1140/epjc/s2005-02169-1. arXiv: hep-ph/0406184 [hep-ph]. Updated results and plots available at: http://ckmfitter.in2p3.fr (zie pag. 199, 201).

Dankwoord

Mijn promotie is een geweldige ervaring geweest, maar was soms ook erg moeilijk. Zonder de hulp en steun van velen was het niet gelukt.

Gerhard en Marcel, bedankt voor jullie begeleiding, advies en vertrouwen. Naast jullie inhoudelijke kennis heb ik vooral tijdens het schrijven van mijn proefschrift heel veel gehad aan jullie ondersteuning.

De geweldige collega's zijn één van de grote voordelen van werken op Nikhef en velen van hen beschouw ik inmiddels als vrienden. Niels, Antonio, Wouter, Veerle, Pieter, Suvayu, Rose, Rosemarie, Siim, Robin, Geert-Jan, Daan, Serena, Chiara en alle anderen, met jullie was er nooit een saai moment tijdens of na het werk.

Lucie, ik kon altijd bij je terecht als ik behoefte had aan een pauze (met een blikje kauwgomballensmaak voor jou), om te klagen, of om gewoon over dingen te praten. Onze tijd in Genève overlapte grotendeels en alle biertjes en etentjes hebben mijn tijd op Cern een stuk leefbaarder gemaakt. Tijdens de master en promotie en ben je mijn beste vriendin geworden; het was een eer dat ik je paranimf mocht zijn en ik kan me zelf geen betere wensen.

Quite soon after the start of my Ph.D, I moved to Geneva to support the LHCb High Level Trigger (HLT) group. I had been doing some HLT related work during the first few months of my Ph.D, but hadn't expected to be working on something as important as pattern recognition for the LHCb muon triggers so soon. Vava, Johannes and the rest of the LHCb trigger group, thank you for your trust in my abilities and making me feel at home at Cern. Working with you has been a pleasure and has opened doors I wouldn't have known existed.

Eric, vanaf het begin van mijn tijd op Cern heb ik met veel plezier met je samengewerkt. Toen ik weer terug in Amsterdam was bood je me aan dat ik bij je kon logeren als ik op Cern moest zijn en door de gezelligheid en de halve grapefruits die elke ochtend voor het ontbijt klaarstonden voelde ik me meteen thuis. Nu mijn hernieuwde verblijf in je zoon's oude slaapkamer al een paar maanden duurt, blijkt je gastvrijheid haast eindeloos te zijn.

Moeders, Steven en Dominique, jullie zijn altijd heel erg trots geweest op waar ik allemaal mee bezig ben. Als wat je doet gewoon wordt, omdat je er dagelijks mee bezig bent, is het heel fijn om eraan herinnerd te worden dat je met iets bijzonders bezig bent.

Lieve Sara, het schrijven van mijn proefschrift viel me zwaar, maar jij hield me op de been en aan de gang. Zonder jouw dagelijkse steun en aanmoediging had ik het niet gered.